Backward Induction Reasoning in Games with Incomplete Information *

Antonio Penta†
University of Wisconsin - Madison,
Dept. of Economics

This version: September 26, 2011

Abstract

Backward Induction is one of the central notions in game theory. A number of solution concepts, such as subgame perfect and sequential equilibrium, are thought of as being based on “backward induction reasoning”. Yet, it is not clear what this means precisely, particularly in situations with incomplete information, where the game cannot even be solved backwards.

This paper introduces a solution concept for games with incomplete information, backward extensive form rationalizability (BR for short), and proves several properties that show how, in a precise sense, BR characterizes the implications of backward induction reasoning in games with incomplete information. These results reconcile in a unitary framework several ideas traditionally (though only informally) associated to the logic of backward induction, such as the idea of deviations as “mistakes”, tremble-based equilibrium concepts, the belief-persistence hypothesis, the notion of subgame consistency and the possibility of solving the game backwards.

JEL Codes: C72; C73; D82.

1 Introduction

Backward induction is one of the basic notions of game theory. Although backward induction is only defined for games with perfect and complete information, the “logic” of backward induction has a much broader scope in game theory. So, for instance, subgame perfect equilibrium is certainly viewed as the natural extension of this logic to games with

---

*Some of the material contained in this manuscript previously circulated as part of my job market paper Robust Dynamic Mechanism Design. I’m indebted to my advisor, George Mailath. I also thank participants of several seminars and conferences.

†email: apenta@ssc.wisc.edu.
imperfect information. But there is a sense in which also solution concepts for incomplete information games, such as sequential or trembling-hand perfect equilibrium, are often thought of as having a backward induction flavor. Yet, it is not clear what “backward induction” means in games with incomplete information. More broadly: What do we mean by “Backward Induction Reasoning”?

Despite its central position in game theory, there is no comprehensive, formal answer to this question. Filling this gap is the aim of this paper.

In pursuit of an answer, a good starting point is to inspect the concepts that are normally associated to the idea of “backward induction reasoning”. Subgame Perfect Equilibrium (SPE) is certainly one of these. SPE is probably the single most common solution concept for dynamic games with complete information. An influential argument in support of SPE is provided by Harsanyi and Selten’s (1988) notion of subgame consistency:

“It is natural to require that a solution function for extensive games is subgame consistent in the sense that the behavior prescribed on a subgame is nothing else than the solution to the subgame” (ibid., p.90)

Subgame consistency warrants SPE the recursive structure of backward induction, i.e. the possibility of determining the solution concept’s predictions for a subgame by looking at the subgame “in isolation”. Hence the possibility (in games with finite horizon) to solve for the subgame perfect equilibria starting from the terminal nodes and proceeding backwards. This is extremely convenient, and certainly one of the main reasons for the prominence of SPE in applied work.

SPE assumes that, even after observing a deviation, agents maintain their beliefs in the equilibrium continuation strategies. This assumption, which can be referred to as the belief persistence hypothesis, often constituted the main target for the numerous critiques to SPE and its implicit assumptions on counterfactuals.1

Several solution concepts extend the ideas of SPE to games with incomplete information, and many of these involve trembles (e.g., trembling-hand perfect equilibrium (Selten, 1975), sequential equilibrium (Kreps and Wilson, 1982), etc.). In these solution concepts, trembles are a shortcut to formalize another idea typically associated to the logic of backward induction: that off-equilibrium moves are “mistakes”, unintended deviations.2

Being the incomplete information counterparts of SPE, it is commonly accepted that these solution concepts share a backward induction flavor. But, if this is true at an intu-

1See e.g. Stalnaker (1996, 1998).

2The view of deviations as “mistakes” contrasts with the logic of forward induction, which requires instead that unexpected moves be rationalized (if possible) as purposeful deviations.
itive level, its precise meaning is not clear, as no formal definition of backward induction is available for games with incomplete information.

One important difference between SPE and its incomplete information counterparts is that the latters lack the recursive structure of SPE. Under incomplete information, an equilibrium requires a specification of agents’ beliefs about the opponents’ types at each information set. But such beliefs are endogenous, equilibrium objects, and must be jointly determined with the equilibrium strategies. With incomplete information, continuation games cannot be considered “in isolation”, and the equilibrium analysis requires the solution of fixed point problems, often difficult to compute. Thus, on the one hand, the “tremble-based” solution concepts supposedly embody the logic of backward induction; but, on the other hand, they lack the recursive structure of SPE, which seems almost a defining feature of backward induction reasoning.

This paper puts forward a solution concept for belief-free dynamic games called Backwards Extensive Form Rationalizability (BR for short). BR consists of an iterated deletion procedure for games in extensive form that at each round eliminates strategies that are not sequential best responses to conjectures that, at each point in the game, must be concentrated on opponents’ continuation strategies that are consistent with the previous rounds of deletion. Through the following results, it is further argued that BR characterizes the behavioral implications of Backward Induction Reasoning in games with incomplete information:

Result 1: BR can be computed by a convenient “backwards procedure” that combines the logic of (normal form) rationalizability and backward induction. The backwards procedure consists of the iterated application of (normal form) rationalizability to the continuation games from each information set considered “in isolation”, starting from the end of the game and then proceeding backwards.

Besides simplifying the computation of the set of BR strategies, this result implies that BR satisfies a property analogous to subgame consistency, which we may call continuation-game consistency: the predictions of BR for each continuation game are nothing but the BR strategies of the continuation game.

I introduce next an equilibrium concept for dynamic Bayesian games, interim perfect equilibrium (IPE). Bayesian games are obtained appending a model of agents’ beliefs, i.e. a type space, to the belief-free game. IPE is the weakest equilibrium notion consistent with sequential rationality and Bayesian updating, and coincides with SPE in complete information games. I show that:

Result 2: IPE is consistent with a tremble-based refinement of Bayesian equilibrium, in which trembles may be correlated with anything.
Result 3: The set of $BR$ strategies in the belief-free game coincides with the set of strategies played as part of some IPE for some type space.

Result 3 says that $BR$ characterizes the “robust predictions” of IPE, that is the IPE predictions that do not depend on assumptions on players’ exogenous beliefs.

Furthermore, while for a given type space the computation of the IPE strategies remains a fixed point problem, Results 1 and 3 together imply that a property analogous to subgame consistency holds for the set IPE strategies: the “robust predictions” are continuation-game consistent.

At a practical level, these results show that rather than computing the set of IPE by solving a large (possibly infinite, in fact) number of fixed point problems, the set of all IPE strategies can be computed by means of a tractable backwards procedure. These results therefore can be particularly useful in applications.\(^3\)

Finally, two epistemic characterizations of $BR$ are provided:

Result 4: $BR$ is characterized by Rationality and the following assumption on agents’ beliefs:

- **Common Certainty of Future Rationality.** That is, players share common certainty of rationality at the beginning of the game. If an unexpected move is observed, players are willing to accept the idea that the deviation was “a mistake” and maintain common certainty of rationality in the continuation game; or

- **Common Certainty of “Full Rationality” and Belief Persistence.** That is, at the beginning of the game, players share common certainty of rationality and they never change their beliefs about the opponents’ continuation strategies. “Full Rationality” is a stronger notion than rationality, in that it refers to statements about agents’ rationality conditional on counterfactual hypothesis. (To accommodate such counterfactual propositions, it will be necessary to innovate on the existing literature by introducing richer epistemic models.)

Overall, the results above (formally) reconcile all the features that are (informally) associated to backward induction reasoning: the recursive structure, the notion of continuation-game consistency, the belief persistence hypothesis, the idea of “trembles” and of deviations as unintended mistakes. There is thus a precise sense in which IPE is the incomplete information counterpart of SPE embodying the backwards induction logic and nothing more.

\(^3\)See Penta (2009) for an application to problems of robust dynamic mechanism design.
I thus argue that the epistemic assumptions of Common Certainty of Future Rationality (or, alternatively, Common Certainty of Rationality and Belief Persistence) provide a comprehensive formal answer to the opening question: “What do we mean by ‘Backward Induction Reasoning’?”

From an applied perspective, this paper provides foundations to a tractable “backwards procedure” that characterizes the set of equilibria of dynamic games with incomplete information. The tractability of the algorithm may prove useful in overcoming the difficulties typically faced in applied and empirical works.

2 Belief-Free Dynamic Games

The analysis that follows concerns multistage games with observable actions. These are defined by an extensive form and agents’ preferences and information.

Extensive Form. The game has $L$ stages, indexed by $l = 1, 2, ..., L$. Let $h^0$ denote the empty history, and for every player $i \in N = \{1,...,n\}$, let $A_i$ denote the (finite) set of actions available to player $i$ throughout the game. At stage $l = 1$, agents $i \in N$ simultaneously choose actions $a_i^1$ from the finite sets $A_i(h^0) \subseteq A_i$ (for each $i \in N$). The chosen action profile is publicly observed, hence the set $H^1 = \times_{i\in N} A_i(h^0)$ denotes the set of histories of length one. For every $h \in H^1$, let $A_i(h) \subseteq A_i$ denote the set of actions available to player $i$ at history $h$. For every $l = 2, ..., L$, the set $H^l$ of public histories of length $l$ is defined recursively as follows: for any $l$ and $h^{l-1} \in H^{l-1}$, let $A_i(h^{l-1}) \subseteq A_i$ denote the set of player $i$’s actions available at history $h^{l-1}$, and let $A_i(h^{l-1}) = \times_{i\in N} A_i(h^{l-1})$ and $A_{-i}(h^{l-1}) = \times_{j\in N\setminus \{i\}} A_j(h^{l-1})$.

$$H^l = \left\{ (h^{l-1}, (a_i^{l})_{i\in N}) \in H^{l-1} \times A_i : a_i^{l} \in A_i(h^{l-1}) \text{ for every } i \in N \right\}.$$  

The set of public histories is defined as $H = \bigcup_{l=0}^{L-1} H^l$ (where $H^0 = \{h^0\}$), while the set of terminal histories is $Z = H^L$.

Without loss of generality, sets $A_i(h)$ are assumed non-empty for each $h \in H$: player $i$ is inactive at $h$ if $|A_i(h)| = 1$; he is active otherwise. This setup allows finitely repeated games as a special case, or games with perfect information if $H$ is such that only one player is active at each $h$. If $H = \{h^0\}$, the game is static. It will be convenient to introduce the precedence relation $\prec$ on $H$: $h^l \prec h^{l+K}$ if and only if there exists $(a_k^k)_{k=1,...,K}$ such that $h^{l+K} = \left( h^l, (a_k^k)_{k=1,...,K} \right)$.

---

4See Fudenberg and Tirole, §3.2 and §8.2. At the expense of heavier notation, the analysis can be easily adapted to all finite dynamic games with perfect recall.

5The set $A_i(h)$ may vary with history $h$, hence the game need not be a repeated game.

6Similar notation will be adopted for other product sets.
Preferences and Information. To model situations with incomplete information, players’ preferences over the terminal nodes are parametrized on a fundamental space of uncertainty 

\[ \Theta = \Theta_0 \times \Theta_1 \times \ldots \times \Theta_n. \]

Elements of \( \Theta \) are referred to as payoff states, and payoff functions are denoted by \( u_i : Z \times \Theta \to \mathbb{R}, \) for each \( i \in N. \) For each \( i = 1, \ldots, n, \) \( \Theta_i \) is the set of player \( i \)’s payoff types. \( \Theta_0 \) is referred to as the set of states of nature. For each \( i, \) \( \Theta_{-i} = \times_{j \in N \setminus \{i\}} \Theta_j, \) so that \( \Theta = \Theta_0 \times \Theta_i \times \Theta_{-i}. \) (To avoid unnecessary technicalities, the set \( \Theta \) is assumed finite throughout.)

When the true state is \( (\theta_0, \theta_1, \ldots, \theta_n), \) player \( i \)’s payoff type is \( \theta_i, \) privately observed at the beginning of the game. Hence, payoff types represent agents’ information about the payoff state: if \( i \)’s payoff type is \( \hat{\theta}_i, \) \( i \) knows that the true state belongs to the set \( \Theta_0 \times \{ \hat{\theta}_i \} \times \Theta_{-i}. \) The set \( \Theta_0 \) represents the residual uncertainty that is left after pooling everybody’s information.

The tuple \( \langle \Theta_0, (\Theta_i, u_i)_{i \in N} \rangle \) thus represents agents’ information about payoffs. It is assumed common knowledge and referred to as preference-information structure (PI-structure). Special cases of interest are: complete information (\( \Theta_k \) is a singleton for all \( k = 0, 1, \ldots, n); \) private values (if, for all \( i \in N, u_i \) is constant in \( (\theta_0, \theta_{-i}); \) no information (if \( u_i \)’s are constant on \( \Theta_{-0}); \) distributed knowledge (if \( u_i \)’s are constant in \( \theta_0). \)

2.1 Belief-Free Games

A belief-free dynamic game is thus defined by a tuple

\[ \Gamma = \langle N, H, Z, \Theta_0, (\Theta_i, u_i)_{i \in N} \rangle. \]

Notice that this is not a Bayesian game, as \( \Gamma \) does not include a model of agents’ interactive beliefs over \( \Theta. \) Bayesian games will be introduced in Section 4.1.

Strategic Forms. Pure strategies in the belief-free game \( \Gamma \) are functions \( s_i : H \to A_i \) such that for each \( h \in H, s_i (h) \in A_i (h). \) The set of player \( i \)’s strategies is denoted by \( S_i, \) and as usual we define the sets \( S = \times_{i \in N} S_i \) and \( S_{-i} = \times_{j \in N \setminus \{i\}} S_j. \) To distinguish them from those that will be introduced for Bayesian games, elements of \( S_i \) are referred to as interim (pure) strategies.

Any strategy profile \( s \in S \) induces a terminal history \( z (\theta) \in Z. \) Hence, we can define strategic-form payoff functions \( U_i : S \times \Theta \to \mathbb{R} \) as \( U_i (s, \theta) = u_i (z (s), \theta) \) for each \( s \) and \( \theta. \)

---

7General information partitions on \( \Theta \) could be considered, at the expense of heavier notation. Restricting attention to product structures entails no essential loss of generality.

8Product structures with distributed knowledge are common in the literature on robust mechanism design. (see, e.g., Bergemann and Morris 2005, 2009)
It’s useful to introduce notation for the *interim mixed strategies* (elements of $\Delta (S_i)$) and the *interim behavior strategies* (elements of $\times_{h_0 \in H} \Delta (A_i)$), which will be used for the analysis of Bayesian games (Section 4.1). To save on notation, we will take advantage of Kuhn’s (1953) equivalence theorem and use the same symbol, $\sigma_i \in \sum_i$, to denote both kinds of strategies (its interpretation will be clear from the context). Notation $\sigma_i [s_i]$ refers to the probability that mixed strategy $\sigma_i$ attaches to strategy $s_i$; while $\sigma_i (a_i|h)$ refers to the probability of action $a_i$ at private history $h$ according to the behavior strategy $\sigma_i$.

For each public history $h$ and player $i$, let $S_i (h)$ denote the set of player $i$’s strategies that are consistent with history $h$ being observed.

It is also convenient to define strategies and payoff functions for the continuations games: For each public history $h \in H$, let $S_i^h$ denote the set of strategies in the continuation game starting from $h$, and for each $s_i \in S_i$, let $s_i|h \in S_i^h$ denote the continuation of $s_i$ from history $h$. The notation $z (s|h, \theta)$ refers to the terminal history induced by strategy profile $s$ from the public history $h$, when the realized state is $\theta$. Strategic-form payoff functions can be defined for continuations from a given public history: for each $h \in H$ and each $(s, \theta) \in S \times \Theta$, let $U_i (s, \theta; h) = u_i (z (s|h), \theta)$. (For the initial history $h^0$, $U_i (s, \theta)$ will be written instead of $U_i (s, \theta; h^0)$.)

### 3 Backwards Extensive Form Rationalizability

Backwards Extensive Form Rationalizability ($BR$) is a solution concept for belief-free games in extensive form, $\Gamma$. Similar to rationalizability, $BR$ is a non-equilibrium solution concept: in game $\Gamma$ agents form conjectures about everyone’s behavior, which may or may not be consistent with each other.

**Endogenous Beliefs: Conjectures.** At every history, players hold *conjectures* about the state of nature, the opponents’ payoff types and (everybody’s) behavior. These are represented by conditional probability systems (CPS), i.e. arrays of conditional beliefs, one for each history. These beliefs differ from those that will be introduced in Section 4.1 in that they concern and depend on endogenous variables such as the opponents’ behavior. Hence, these are *endogenous beliefs*. To avoid confusion, we thus refer to this kind of beliefs as “conjectures”, retaining the term “beliefs” for those introduced in Section 4.1.

For each history $h \in H$, define the event $[h]_i \subseteq \Theta_0 \times \Theta_{-i} \times S$ as:

$$[h]_i = \Theta_0 \times \Theta_{-i} \times S (h).$$

(Notice that, by definition, $[h]_i \subseteq [h']_i$ whenever $h$ follows $h'$.)

**Definition 1** A conjecture for agent $i$ is a conditional probability system (CPS hereafter), that is a collection $\mu^i = (\mu^i (h))_{h \in H}$ of conditional distributions $\mu^i (h) \in \Delta (\Theta_0 \times \Theta_{-i} \times S)$
that satisfy the following conditions:

C.1 For all \( h \in \mathcal{H} \), \( \text{supp}(\mu^i(h)) \subseteq [h]_i \);

C.2 For every measurable \( A \subseteq [h]_i \subseteq [h']_i \), \( \mu^i(h)[A] \cdot \mu^i(h')[h]_i = \mu^i(h')[A] \).

The set of CPS over \( \Theta_0 \times \Theta_{-i} \times S \) is denoted by \( \Delta^H(\Theta_0 \times \Theta_{-i} \times S) \).

Condition C.1 states that agents’ are always certain of what they know; condition C.2 states that agents’ conjectures are consistent with Bayesian updating whenever possible. Notice that in this specification agents entertain conjectures about the payoff state, the opponents’ and their own strategies. The latter point is not entirely standard: in similar non-equilibrium solution concepts for games in extensive form it’s common practice to model conjectures on the opponents’ behavior only (see, e.g. Battigalli and Siniscalchi, 2007, or Penta, 2010). We will discuss this point in some detail in Section 4.

**Sequential Rationality.** Strategy \( s_i \) is sequentially rational for type \( \theta_i \) with respect to conjectures \( \mu^i \) if, at each history \( h \in \mathcal{H} \), it prescribes optimal behavior in the continuation game with respect to conjectures \( \mu^i(h) \). Formally: for any \( \hat{\theta}_i \in \Theta_i \), given a CPS \( \mu^i \in \Delta^H(\Theta_0 \times \Theta_{-i} \times S) \) and a history \( h \), \( \hat{\theta}_i \)'s expected payoff from \( s_i \) at \( h \), given \( \mu^i \), is defined as:

\[
U_i(s_i, \mu^i; h, \hat{\theta}_i) = \sum_{\theta_0, \theta_{-i}, s_{-i}} \text{marg}_{\Theta_0 \times \Theta_{-i} \times S_{-i}} \mu^i(h)[\theta_0, \theta_{-i}, s_{-i}] \cdot U_i(s_i, s_{-i}, \theta_0, \hat{\theta}_i, \theta_{-i}; h)
\]

**Definition 2** Strategy \( s_i \) is sequentially rational for payoff-type \( \theta_i \) with respect to \( \mu^i \in \Delta^H(\Theta_0 \times \Theta_{-i} \times S) \), written \( s_i \in r_i(\mu^i, \theta_i) \), if and only if for each \( h \in \mathcal{H} \) and each \( s'_i \in S_i \) the following inequality is satisfied:

\[
\bar{U}_i(s_i, \mu^i; h, \theta_i) \geq \bar{U}_i(s'_i, \mu^i; h, \theta_i).
\]

If \( s_i \in r_i(\mu^i, \theta_i) \), we say that conjectures \( \mu^i \) “justify” strategy \( s_i \) for type \( \theta_i \).

### 3.1 Backwards Rationalizability in the Extensive Form

We introduce next the solution concept that, it will be argued, characterizes backward induction reasoning in incomplete information games: Backwards Extensive Form Rationalizability (B\( \mathcal{R} \)).
Definition 3 For each \( i \in N \) let \( \mathcal{BR}^0_i = \Theta_i \times S_i \). Recursively, for \( k = 1, 2, \ldots \), let \( \mathcal{BR}^{k-1}_i = \times_{j \in N \setminus \{i\}} \mathcal{BR}^{k-1}_j \), and for each \( \theta_i \in \Theta_i \), let,

\[
\mathcal{BR}^k_i (\theta_i) = \left\{ \begin{array}{l}
\exists \mu^i \in \Delta^N (\Theta_0 \times \Theta_{-i} \times S) \text{ s.t.} \\
(1) \hat{s}_i \in r_i (\mu^i, \theta_i) \\
(2) \text{supp} (\mu^i (h^0)) \subseteq \Theta_0 \times \mathcal{BR}^{k-1}_{-i} \times \{\hat{s}_i\} \\
(3) \text{for each } h \in H: \\
(3.1) s_i|h = \hat{s}_i|h, \text{ and} \\
(3.2) \exists s'_{-i} \in \mathcal{BR}^{k-1}_{-i} (\theta_{-i}) : s'_{-i}|h = s_{-i}|h \n\end{array} \right\},
\]

\[
\mathcal{BR}^k_i = \{ (\theta_i, s_i) \in \Theta_i \times S_i : s_i \in \mathcal{BR}^k_i (\theta_i) \}, \mathcal{BR}^k = \times_{i \in N} \mathcal{BR}^k_i, \text{ and finally } \mathcal{BR} := \bigcap_{k \geq 0} \mathcal{BR}^k.
\]

\( \mathcal{BR} \) consists of an iterated deletion procedure. At each round, strategy \( \hat{s}_i \) survives for type \( \theta_i \) if it is justified by conjectures \( \mu^i \) that satisfy two conditions: condition (2) states that at the beginning of the game, the agent must be certain of his own strategy \( \hat{s}_i \) and have conjectures concentrated on the opponents’ type and strategies consistent with the previous rounds of deletion; condition (3) restricts the agent’s conjectures at unexpected histories: condition (3.1) states that agent \( i \) is always certain of his own continuation strategy; condition (3.2) requires conjectures to be concentrated on opponents’ continuation strategies that are consistent with the previous rounds of deletion. Notice however that agents’ conjectures about \( \Theta_0 \times \Theta_{-i} \) at unexpected histories are unrestricted. Thus, condition (3) embeds two conceptually distinct kinds of assumptions: the first concerning agents’ conjectures about \( \Theta_0 \times \Theta_{-i} \); the second concerning their conjectures about the continuation behavior. For ease of reference, they are summarized as follows:

- **Unrestricted-Inference Assumption (UIA):** At unexpected histories, agents’ conjectures about \( \Theta_0 \times \Theta_{-i} \) are unrestricted. In particular, agents are free to infer anything about the opponents’ private information from the public history.

For example, conditional conjectures may be such that \( \text{marg}_{\Theta_{-i}} \mu^i (h) \) is concentrated on opponents’ types \( \theta_{-i} \) for whom some of the previous moves in \( h \) would be irrational, or “mistakes”. Nonetheless, condition (3.2) implies that it is believed that such types \( \theta_{-i} \) will behave rationally in the future. From an epistemic viewpoint, it can be shown that \( \mathcal{BR} \) can be interpreted as common certainty of future rationality at every history (for the formal statement, see Section 6.3)

- **Common Certainty in Future Rationality (CCFR):** at every history (expected or not), agents share common certainty in future rationality.
CCFR can be interpreted as a condition of belief persistence on the continuation strategies. (A formal connection between $\mathcal{BR}$ and the belief persistence hypothesis is provided in Section 6.4).

**Remark 1** Since the game is finite, the iterated deletion procedure stops after finitely many rounds: there exists $K < \infty$ such that $\mathcal{BR}^K = \mathcal{BR}^{K+1}$. It’s also trivial to show that $\mathcal{BR}$ satisfies the following fixed point characterization:

\begin{align*}
\text{Lemma 1} \quad \text{Strategy } s_i \in \mathcal{BR}_i(\theta_1) \text{ if and only if } & \exists \mu^i \in \Delta^{\mathcal{H}_i}(\Theta_0 \times \Theta_1 \times S) \text{ s.t. } (1) \ s_i \in r_i(\mu^i, \theta_1); (2) \ \text{supp}(\mu^i(h^0)) \subseteq \Theta_0 \times \mathcal{BR}_{-i} \times \{s_i\} \text{ and (3) for each } h \in \mathcal{H}: s \in \text{supp}(\text{marg}_s \mu^i(h)) \\
& \text{implies: (3.1) } s_i|h = s_i|h, \text{ and (3.2) } \exists (\theta_{-i}, s'_{-i}) \in \mathcal{BR}_{-i}: s'_{-i}|h = s_{-i}|h.
\end{align*}

**Example 1** Consider the game in Figure 1, and let $\Theta_1 = \{10, -10\}$, while $\Theta_0$ and $\Theta_2$ are singletons (hence, player 1 knows the true state, while player 2 has no information). Let’s apply $\mathcal{BR}$: at the first round, strategies involving $L_3$ and $R_3$ are deleted for all types of player 1, while strategies involving $l$ (resp. $r$) are deleted for type $\theta_1 = -10$ (resp. $\theta_1 = 10$). So, for instance, after the first round strategy $rL_1R_2$ survives for type $\theta_1 = 10$. Now, suppose that 2’s initial conjectures are concentrated on payoff state-strategy pair $(\theta_1, s_1) = (-10, rL_1R_2)$: then, $l$ is unexpected, so after observing it we are free to specify 2’s beliefs, who could for example assign probability one on pair $(\theta_1, s_1) = (10, lL_1R_2)$, hence play $a_2$, or on pair $(\theta_1, s_1) = (-10, lL_1R_2)$, hence play $a_1$. Similarly, both $b_1$ and $b_2$ in the right-most continuation game can be justified if 2’s initial conjectures assign probability one to $\theta_1 = 10$. Hence, strategies that survive $\mathcal{BR}$ are $\{a_1b_2, a_1b_3, a_2b_2, a_2b_3\}$ for player 2, $\{lL_1R_1, lL_1R_2, lL_2R_1, lL_2R_2\}$ for type $\theta_1 = 10$ and $\{rL_1R_1, rL_1R_2, rL_2R_1, rL_2R_2\}$ for type $\theta_1 = -10$.  

![Figure 1: Example 1](image-url)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>1</th>
<th>2</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_1$</td>
<td>2</td>
<td>$1$</td>
<td>$1 + \theta_1$</td>
<td>0</td>
<td>$-1$</td>
<td>$R_1$</td>
<td>2</td>
<td>$-\theta_1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$L_2$</td>
<td>0</td>
<td>$1$</td>
<td>$2 + \theta_1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$R_2$</td>
<td>$-\theta_1$</td>
<td>$1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$L_3$</td>
<td>$-1$</td>
<td>$2$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$R_3$</td>
<td>$-1$</td>
<td>$-\theta_1$</td>
<td>$3$</td>
</tr>
</tbody>
</table>
4 Backwards Rationalizability and Equilibria

This section explores the connection between B&R and equilibrium predictions. Pursuing an equilibrium approach in games with incomplete information requires a specification of agents’ hierarchies of beliefs. We will follow the traditional approach (Harsanyi, 1967-68) of modeling such hierarchies of beliefs implicitly, by means of type spaces. Appending a type space to a belief-free game delivers a Bayesian game.

4.1 Bayesian Games

Definition 4 A (Θ-based) type space is a tuple

\[ T = (T_i, \theta_i, \tau_i)_{i \in N} \]

such that for each \( i \in N \), \( T_i \) is a finite set of types, \( \theta_i : T_i \to \Theta_i \) is an onto function assigning a payoff-type to each type, and \( \tau_i : T_i \to \Delta (\Theta \times T_{-i}) \) assigns to each type a belief about the payoff state and the opponents’ types.\(^9\)

The Bayesian game obtained appending type space \( T = (T_i, \theta_i, \tau_i)_{i \in N} \) to the belief-free game \( \Gamma \) is defined as the tuple

\[ \Gamma^T = \langle N, \Theta, \mathcal{H}, \mathcal{Z}, (T_i, \tau_i, \theta_i, \hat{u}_i)_{i \in N} \rangle \]

where \( \hat{u}_i : \mathcal{Z} \times \Theta_0 \times T \to \mathbb{R} \) is such that that for each \( (z, \theta_0, t) \in \mathcal{Z} \times \Theta_0 \times T \), \( \hat{u}_i (z, \theta_0, t) = u_i (z, \theta_0, \theta (t)) \). To avoid unnecessary notation, in the following we will use \( u_i \) to denote both payoff functions.

Strategies in a Bayesian game are functions \( b_i : T_i \to \Sigma_i \) assigning an interim (mixed or behavior) strategy to each type in the type space. The notation \( b_i (a_i^t; t_i, h^{l-1}) \) refers to the probability that behavior strategy \( b_i (t_i) \in \Sigma_i \) assigns to action \( a_i^t \) at \( h^{l-1} \).

4.2 Interim Perfect Equilibrium

Define the set of information sets of player \( i \) in the Bayesian game as \( \mathcal{H}_i = T_i \times \mathcal{H} \).\(^{10}\) A system of beliefs consists of collections \( (p_i (h_i))_{h_i \in \mathcal{H}_i} \) for each agent \( i \), such that \( p_i (h_i) \in \Delta (\Theta_0 \times T_{-i}) \) for each \( h_i \in \mathcal{H}_i \): a belief system represents agents’ conditional beliefs about the state of nature and the opponents’ types at each information set. A strategy profile and a belief system \( (b, p) \) form an assessment. For each agent \( i \), a strategy profile \( b \)

\[^9\]The restriction to finite type spaces is only made for simplicity of exposition.

\[^{10}\]Behavior strategies in the Bayesian game could be defined as functions \( \gamma_i : \mathcal{H}_i \to A_i \) s.t. \( \gamma_i (t_i, h) \in \Delta (A_i (h)) \) for each \( h \in \mathcal{H} \). Clearly, this is equivalent to the definition of strategies as functions \( b_i : T_i \to \Sigma_i \).

11
and conditional beliefs \( p_i \) induce, at each private history \( h_i^{l-1} = (t_i, h_i^{l-1}) \), a probability distribution \( P^{\sigma, p_i} (a_i | h_i^{l-1}) \) over action profiles at stage \( l \):

\[
P^{b, p_i} (a_i; h_i^{l-1}) = b_i (a_i; h_i^{l-1}) \cdot \sum_{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}} p_i (\theta_0, t_{-i} | h_i^{l-1}) \cdot \prod_{j \neq i} b_j (a_j; t_j, h_i^{l-1})
\]

**Definition 5** Assessment \((b, p)\) is weakly preconsistent if for each \( i \in N \):

\[
\forall t_i \in T_i, p_i (t_i, h_0) = \tau_i (t_i)
\]

\[
\forall h_i^l = (h_i^{l-1}, a_i) \in H_i,
\quad p_i (\theta_0, t_{-i} | h_i^l) = p_i (\theta_0, t_{-i} | h_i^{l-1}) \cdot P^{b, p_i} (a_i; h_i^{l-1}).
\]

Condition (4) requires each agent’s beliefs conditional on observing type \( t_i \) to agree with that type’s (exogenous) beliefs as specified in the type space \( T \); condition (5) requires that the belief system \( p_i \) is consistent with Bayesian updating whenever possible.

From the point of view of each \( i \), for each \( h_i = (t_i, h) \in H_i \) and strategy profile \( b \), the induced terminal history is a random variable that depends on the opponents’ type profile. This is denoted by \( z (b | h_i, t_{-i}) \). As done for belief-free games (Section 2), we can define strategic-form payoff functions for the continuation games:

\[
\hat{U}_i (b, \theta_0, t_i, t_{-i}; h_i) = u_i (z (b | h_i, t_{-i}), \theta (t)).
\]

**Definition 6** Fix a belief system \( p \). Strategy profile is sequentially rational with respect to \( p \) if for every \( i \in N \) and every \( h_i^l \in H_i \setminus \{ \phi \} \), the following inequality is satisfied for every \( b_i^l : T_i \rightarrow \Sigma_i \):

\[
\sum_{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}} \hat{U}_i (b, \theta_0, t_i, t_{-i}; h_i^l) \cdot p_i (\theta_0, t_{-i} | h_i^l) \\
\geq \sum_{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}} \hat{U}_i (b_i^l, \theta_0, t_i, t_{-i}; h_i^l) \cdot p_i (\theta_0, t_{-i} | h_i^l).
\]

**Definition 7** An assessment \((\sigma, p)\) is an Interim Perfect Equilibrium (IPE) if it is weakly preconsistent and sequentially rational.

If inequality (6) is only imposed at private histories of length zero, the solution concept coincides with interim equilibrium (Bergemann and Morris, 2005). IPE refines interim equilibrium imposing two natural conditions: first, sequential rationality; second, preconsistency of the belief system.
Notice that *weak preconsistency* imposes no restrictions on the beliefs held at histories that receive zero probability at the preceding node.\(^{11}\) Hence, even if agents’ initial beliefs admit a common prior, IPE is weaker than Fudenberg and Tirole’s (1991) perfect Bayesian equilibrium. Also, notice that any player’s deviation is a zero probability event, and treated the same way. In particular, if history \(h_i^t\) is precluded by \(b_i(h_i^{t-1})\) alone, \(h_i^t \notin \text{supp} P_{\sigma_i,\rho_i}(h_i^{t-1})\), and agent \(i\)'s beliefs at \(h_i^t\) are unrestricted the same way they would be after an unexpected move of the opponents. This feature of IPE is not entirely standard, but it is key to the result that the set of IPE strategies (taking the union over all type spaces) can be computed by means of a convenient “backwards procedure”: Treating own deviations the same as the opponents’ is key to the possibility of considering continuation games “in isolation”, necessary for the result.\(^{12}\)

**Tremble-Based Formulation.** It can be shown that IPE is consistent with a “trembling-hand” view of unexpected moves, in which no restrictions on the possible correlations between “trembles” and other elements of uncertainty are imposed.

### 4.3 Characterization of the set of IPE

As emphasized above, in \(\mathcal{BR}\) agents hold conjectures about both the opponents’ and their own strategies. First, notice that conditions (2) and (3.2) in the definition of \(\mathcal{BR}\) maintain that agents are always certain of their own strategy; furthermore, the definition of sequential best response (def. 2) depends only on the marginals of the conditional conjectures over \(\Theta_0 \times \Theta_{-i} \times S_{-i}\). Hence, this particular feature of \(\mathcal{BR}\) does not affect the standard notion of rationality. The fact that conjectures are elements of \(\Delta^H(\Theta_0 \times \Theta_{-i} \times S)\) rather than \(\Delta^H(\Theta_0 \times \Theta_{-i} \times S_{-i})\) corresponds to the assumption that IPE treats all deviations the same; its implication is that both histories arising from unexpected moves of the opponents and from one’s own deviations represent zero-probability events, allowing the same set of conditional beliefs about \(\Theta_0 \times \Theta_{-i} \times S\), with essentially the same freedom that IPE allows after anyone’s deviation. This is the main insight behind the following result (the proof is in Appendix A.1).

\(^{11}\)Unlike other notions of weak perfect Bayesian equilibrium, in IPE agents’ beliefs are consistent with Bayesian updating also off-the-equilibrium path. In particular, in complete information games, IPE coincides with subgame-perfect equilibrium.

\(^{12}\)In Penta (2009b) I consider a minimal strengthening of IPE, in which agents’ beliefs are not upset by unilateral own deviations, and I show how the analysis that follows adapts to that case: The “backwards procedure” to compute the set of equilibria across models of beliefs must be modified, so to keep track of the restrictions the extensive form imposes on the agents’ beliefs at unexpected nodes. The possibility of envisioning continuation games “in isolation” is thus lost.
Proposition 1 Fix a belief-free game $\Gamma$. For each $i$, $s_i \in BR_i(\hat{\theta}_i)$ if and only if $\exists T$, $\hat{i}_i \in T_i$ and $(\hat{b}, \hat{p})$ such that: (i) $(\hat{b}, \hat{p})$ is an IPE of $\Gamma^T$; (ii) $\theta_i(\hat{i}_i) = \hat{\theta}_i$ and (ii) $\hat{s}_i \in \text{supp} \hat{b}_i(\hat{i}_i)$. An analogous result can be obtained for the more standard refinement of IPE, in which unilateral own deviations leave an agents’ beliefs unchanged, applying to a modified version of $BR$: such modification entails assuming that agents only form conjectures about $\Theta_0 \times \Theta_{-i} \times S_{-i}$ (that is, $\mu^i \in \Delta^H(\Theta_0 \times \Theta_{-i} \times S_{-i})$) and by consequently adapting conditions (2) and (3) in the definition of $BR$. (See Penta, 2009b.) Hence, the assumption that IPE treats anyone’s deviation the same (and, correspondingly, that in $BR$ agents hold conjectures about their own strategy as well) is not crucial to characterize the set of equilibrium strategies across models of beliefs. It is crucial instead for the next result, which shows that such set can be computed applying a procedure that extends the logic of backward induction to environments with incomplete information (proposition 2 below).

5 Continuation-game Consistency

This Section provides a characterization of $BR$ in terms of a recursive procedure, that solves the game backwards. Together with proposition 1, the result in this section implies that the robust predictions of IPE satisfy a property analogous to Harsanyi and Selten’s (1988) subgame consistency: for each $h$, the set of IPE continuation strategies from $h$ coincides with the set of IPE strategies of the continuation game considered in isolation.

The Backwards Procedure. The backwards procedure is described as follows: Fix a public history $h^{L-1}$ of length $L - 1$. For each payoff-type $\theta_i \in \Theta_i$ of each agent, the continuation game is a static game, to which we can apply belief-free rationalizability (e.g., Bergemann and Morris, 2009). For each $h^{L-1}$, let $R_i^{h^{L-1}}$ denote the set of pairs $(\theta_i, s_i|h^{L-1})$ such that continuation strategy $s_i|h^{L-1}$ is rationalizable in the continuation game from $h^{L-1}$ for type $\theta_i$. We now proceed backwards: for each public history $h^{L-2}$ of length $L - 2$, we apply again rationalizability to the continuation game from $h^{L-2}$ (in normal form), restricting continuation strategies $s_i|h^{L-2} \in S_i^{h^{L-2}}$ to be rationalizable in the continuation games from histories of length $h^{L-1}$. $R_i^{h^{L-2}}$ denotes the set of pairs $(\theta_i, s_i|h^{L-2})$ such that continuation strategy $s_i|h^{L-2}$ is rationalizable in the continuation game from $h^{L-2}$ for type $\theta_i$. Inductively, this is done for each $h^{l-1}$, $l = L, \ldots, 1$, until the initial node is reached.

Before introducing the procedure formally, consider the following example:

Example 2 Consider the game in figure 1 again. If we apply belief-free rationalizability to the continuation game following $r$, $R_3$ is deleted at the first round for both types of
player 1, and \( b_1 \) at the second round for player 2. In this continuation game, the procedure selects continuations \( \{ b_2, b_3 \} \) for player 2 and continuations \( \{ R_1, R_2 \} \) for both types of player 1. Similarly, after \( l \), belief-free rationalizability selects \( \{ a_1, a_2 \} \) for player 2, and \( \{ L_1, L_2 \} \) for both types of player 1. So, now we apply belief-free rationalizability to the normal form in which it is maintained that continuation strategies are rationalizable in the corresponding continuations, i.e. the relevant strategy sets now are \( \{ a_1 b_2, a_1 b_3, a_2 b_2, a_2 b_3 \} \) for player 2, and

\[
\{ l L_1 R_1, l L_1 R_2, l L_2 R_1, l L_2 R_2 \} \cup \{ r L_1 R_1, r L_1 R_2, r L_2 R_1, r L_2 R_2 \}
\]

for (both types of) player 1: at this stage, type \( \theta_1 = 10 \) deletes all strategies involving \( r \) at the first round, and so does type \( \theta_1 = -10 \) with those involving \( l \), but player 2 doesn’t delete anything: if he expects \( \theta_1 = 10 \) (hence \( 1 \) to play \( l \)), then both \( a_2 b_2 \) and \( a_2 b_3 \) are best responses in the normal form; similarly, if he expects \( \theta_1 = 10 \) (i.e. 1 to play \( r \)), then both \( a_2 b_3 \) and \( a_1 b_3 \) are optimal. Notice that the resulting strategies are precisely those selected by \( BR \) in example 1. As proposition 2 shows, this insight has general validity.\( \diamond \)

The backwards procedure is defined recursively, starting from the last stage of the game and proceeding backwards:

- [\( l = L - 1 \)] For each \( h^{L-1} \in H^{L-1} \), and for each \( \theta_i \in \Theta_i \), let \( R_i^0 (\theta_i, h^{L-1}) = S_i^{h^{L-1}} \)
  For each \( k = 1, 2, ..., \), let

\[
R_i^{k-1} (h^{L-1}) = \left\{ \left( \theta_i, s_i^{h^{L-1}} \right) : s_i^{h^{L-1}} \in R_i^{k-1} (\theta_i, h^{L-1}) \right\},
\]

\[
R_i^k (\theta_i, h^{L-1}) = \left\{ \begin{array}{l}
\left( \theta_i, s_i^{h^{L-1}} \right) : s_i^{h^{L-1}} \in R_i^{k-1} (\theta_i, h^{L-1}) \\
\left( R.1 \right): \exists \pi \in \Delta (\Theta_0 \times R_i^{k-1} (h^{L-1})) \\
\left( R.2 \right): \text{for all } s_i' \in S_i^{h^{L-1}}, \\
\sum_{\theta_0, \theta_{-i}} U_i \left( s_i^{h^{L-1}}, s_{-i}^{h^{L-1}}, \theta_i, \theta_{-i}; h^{L-1} \right) \\
- U_i \left( s_i' s_{-i}^{h^{L-1}}, \theta_i, \theta_{-i}; h^{L-1} \right) \geq 0
\end{array} \right. 
\]

\[
R_i (h^{L-1}) = \bigcap_{k=1}^{\infty} R_i^k (h^{L-1})
\]
For each $h^l \in H^l$, and for each $\theta_i \in \Theta_i$, let

$$R_0^i(\theta_i, h^l) = \left\{ s_i^{h^l} \in S_i^{h^l} : \forall a^{l+1} \in A^{l+1}(h^l), s_i^{h^l}|(h^l, a^{l+1}) \in R_i(\theta_i, (h^l, a^{l+1})) \right\}.$$  

For each $k = 1, 2, ..., \exists \pi \in \Delta(\Theta_0 \times R^{-1}_{-i}(h^l))$

$$R_k^i(\theta_i, h^l) = \left\{ s_i^{h^l} \in R_{k-1}^i(\theta_i, h^l) : \sum_{\theta_0, \theta_-i, s_-i} \left[ U_i(s_i^{h^l}, s_-i^l, \theta_i, \theta_-i; h^l) - U_i(s_i^{h^l}, s_-i^l, \theta_i, \theta_-i; h^l) \right] \cdot \pi(\theta_0, \theta_-i, s_-i^l) \geq 0 \right\}.$$  

$$R_i(h^l) = \bigcap_{k=1}^{\infty} R_k^i(h^l)$$

**Proposition 2** $BR_i = R_i(h^0)$ for each $i$.

The properties UIA and CCFR discussed in Section 3 provide the basic insight behind this result. First, notice that under UIA, the set of beliefs agents are allowed to entertain about the state of nature and the opponents’ payoff-types is the same at every history ($\Theta_0 \times \Theta_-i$). Hence, in this respect, their information about the opponents’ types in the continuation game from history $h$ is the same as if the game started from $h$. Also, CCFR implies that agents’ assumptions about everyone’s behavior in the continuation is also the same at every history. Thus, UIA and CCFR combined imply that, from the point of view of $BR$, a continuation from history $h$ is equivalent to a game with the same space of uncertainty and strategy spaces equal to the continuation strategies, which justifies the possibility of analyzing continuation games “in isolation”.

### 6 Epistemic Characterizations

This Section provides two alternative epistemic characterizations of $BR$, one in terms of **Common Certainty of Future Rationality**, the other in terms of **Common Certainty of “Full” Rationality and in Belief Persistence**.

The epistemic models adopted here have one important non-standard feature. In existing epistemic models (e.g., Battigalli and Siniscalchi 2002, 2007), a state of the world is a tuple $\omega = (\theta_0, (\theta_i, s_i, \psi_i)_{i \in N})$, i.e. a description of a state of nature ($\theta_0$), and for each player $i$, his payoff type $\theta_i$, strategy $s_i$ and epistemic type $\psi_i$. Each epistemic type induces a CPS over the states of world. In the epistemic models considered here, a state of the world is of the form $\omega = (\theta_0, (\theta_i, x_i, \psi_i)_{i \in N})$, where $x_i : \Theta_i \rightarrow S_i$ denotes player $i$’s *ex-ante strategy*. Ex-ante strategies represent a “full theory” of $i$’s behavior:
in state $\omega = (\theta_i^\omega, (\theta_i^\omega, x_i^\omega, \psi_i^\omega)_{i \in N}) \in \Omega$, player $i$’s “actual disposition to act” is given by strategy $x_i^\omega(\theta_i^\omega)$, but $x_i^\omega$ also represents player $i$’s disposition to act under the hypothesis (counterfactual at $\omega$) that his payoff type is different from $\theta_i^\omega$. Epistemic types $\psi_i$ still induce a CPS over the states of the world $\omega$, but this means that now they express richer information. For instance, consider a state at which $i$’s beliefs are concentrated on a pair $\omega_i$. This means that $i$ believes that the opponents’ type is $\omega_i$, and that his “actual” strategy will be $x_i(\omega_i)$. But it also expresses $i$’s view of how the opponents’ behavior would be under the hypothesis that $\theta_i \neq \theta_i$. This “enrichment” of the state space allows to disentangle epistemic counterfactuals concerning the opponents’ behavior from counterfactuals concerning the opponents’ private information, necessary for the characterization of $\mathcal{BR}$ in terms of Common Certainty of “Full” Rationality and Belief Persistence.\footnote{In environments with complete information, Stalnaker (1996, 1998) discusses the necessity of incorporating different kinds of counterfactual propositions for the analysis of backward induction reasoning. The introduction of ex-ante strategies in the definition of the states of the world is one way of accommodating the richer set of counterfactuals required by the presence of incomplete information.}

The characterization in terms of Common Certainty of Future Rationality doesn’t exploit the information contained in such counterfactual statements.

The epistemic characterizations below will be of the form

$$\mathcal{BR} = \{ (\theta_i, s_i) : \exists \omega \in E \text{ s.t. } (\theta_i^\omega, x_i^\omega(\theta_i^\omega)) = (\theta_i, s_i) \},$$  

for some measurable event $E \subseteq \Omega$. Since each $X_i$ can be seen as a subset of $\Theta_i \times S_i$, equation (7) is written for short as $\mathcal{BR} = \text{proj}_{\Theta_i \times S_i} E$.

### 6.1 Conditional Probability Systems

Let $\Omega$ be a metric space and $\mathcal{A}$ its Borel sigma-algebra. Fix a non-empty collection of subsets $\mathcal{C} \subseteq \mathcal{A} \setminus \emptyset$, to be interpreted as “relevant hypothesis”. A conditional probability system (CPS hereafter) on $(\Omega, \mathcal{A}, \mathcal{C})$ is a mapping $\mu : \mathcal{A} \times \mathcal{C} \to [0, 1]$ such that:

**Axiom 1** For all $B \in \mathcal{C}$, $\mu (B) [B] = 1$

**Axiom 2** For all $B \in \mathcal{C}$, $\mu (B)$ is a probability measure on $(\Omega, \mathcal{A})$.

**Axiom 3** For all $A \in \mathcal{A}$, $B, C \in \mathcal{C}$, if $A \subseteq B \subseteq C$ then $\mu (B) [A] \cdot \mu (C) [B] = \mu (C) [A]$.

The set of CPS on $(\Omega, \mathcal{A}, \mathcal{C})$, denoted by $\Delta^\mathcal{C}(\Omega)$, can be seen as a subset of $[\Delta (\Omega)]^\mathcal{C}$ (i.e. mappings from $\mathcal{C}$ to probability measures over $(\Omega, \mathcal{A})$). CPS’s will be written as $\mu = (\mu (B))_{B \in \mathcal{C}} \in \Delta^\mathcal{C}(\Omega)$. The subsets of $\Omega$ in $\mathcal{C}$ are the conditioning events, each inducing beliefs over $\Omega$; $\Delta (\Omega)$ is endowed with the topology of weak convergence of measures and
[Δ(Ω)]^C is endowed with the product topology. In the belief-free games of Section 2.1, player i’s CPS were obtained setting Ω = Θ_0 × Θ_{-i} × S, and the set of conditioning events was naturally provided by the set of histories \( H \): for each public history \( h \in H \), the corresponding event \([h]_i \) was defined as \([h]_i = Θ_0 × Θ_{-i} × S(h)\).

### 6.2 Epistemic Models

Ex ante strategies assign an interim strategy to every payoff type. The set of ex ante strategies is denoted by \( X_i = S_i^{E_i} \). For each \( x_i \in X_i \), let \( x_i^h \) denote the continuation strategy \( x_i^h : Θ_i → S_i^h \) s.t. for each \( θ_i ∈ Θ_i \), \( x_i^h(θ_i) = x_i(θ_i) | h \).

**Definition 8** A (dynamic) type space on \( Γ \) is a tuple \( \mathcal{D} = \langle N, H, (Θ_i, X_i, Ψ_i, Ω_i, g_i)_{i∈N} \rangle \) s.t. for every \( i ∈ N \), \( Ψ_i \) is a Polish space, and

1. \( Ω_i \subseteq Θ_i × X_i × Ψ_i \) is a closed subset such that \( \text{proj}_{Θ_i × X_i} Ω_i = Θ_i × X_i \)

2. \( g_i : Ψ_i → Δ^H(Θ_0 × Ω_{-i}) \) is a continuous mapping.

For any \( i ∈ N \), the elements of the set \( Ψ_i \) are referred to as Player i’s epistemic types. A type space is compact if all the sets \( (Ψ_i)_{i ∈ I} \) are compact. Let \( Ω = Θ_0 × Ω_1 × ... × Ω_n \) denote the set of states of the world, and let \( A \) be its Borel sigma-algebra.

Thus, at any “possible world” \( ω ∈ Ω = Θ_0 × Ω_1 × ... × Ω_n \), we specify a payoff state \( (θ_0^ω, θ_1^ω, ..., θ_n^ω) \) as well as each player i’s dispositions to act (his strategy \( s_i^ω = x_i^ω(θ_i^ω) \)) and his dispositions to believe (his CPS \( g_i(ψ_i^ω) = (g_i(h(ψ_i^ω)))_{h∈H} \)). As discussed above, \( x_i^ω \) also specifies i’s beliefs about his own strategy under the counterfactual (at \( ω \)) hypothesis that \( θ_i' ≠ θ_i^ω \). Furthermore, notice that these dispositions also include what a player would do and think at histories that are inconsistent with \( ω \) (i.e. \( h ∉ H(s^ω) \)).

As standard in interactive epistemology literature, we assume that players also have beliefs about their own strategy, and complete a player’s system of conditional beliefs by assuming that he is certain of his true ex ante strategy, information type and epistemic type. More specifically, we assume that for every state of the world \( ω \) and every history \( h \), player \( i ∈ N \) would be certain of \( ψ_i^ω \) and \( θ_i^ω \) given \( h \), and would also be certain of \( x_i^ω \) given \( h \), provided that \( x_i^ω(θ_i^ω) ∈ S_i(h) \). We also assume that if \( x_i^ω(θ_i^ω) ∉ S_i(h) \), player \( i \) would still be certain that his continuation strategy agrees with \( x_i^ω(θ) | h \) for each \( θ \). Formally, player i’s conditional beliefs on \( (Ω, H) \) are given by a continuous mapping

\[ g_i^* = (g_i^* h)_{h∈H} : Ω_i → Δ^H(Ω) \]

derived from \( g_i \) by the following formula: for all \((θ_i, x_i, ψ_i) ∈ Ω_i, \) for all \( h ∈ H \) and \( E ∈ A \),

\[ g_i^* h(θ_i, x_i, ψ_i) = g_i(h)(ψ_i) \{ (θ_0, ω_i) ∈ Θ_0 × Ω_{-i} : (θ_0, ω_i, (θ_i, x_i', ψ_i)) ∈ E, \text{ for some } x_i' ∈ X_i \text{ s.t. } x_i'(θ_i') | h = x_i(θ_i') | h \text{ for all } θ_i' ∈ Θ_i \} \]
Type spaces encode a collection of infinite hierarchies of CPSs for each epistemic type of each player. Battigalli and Siniscalchi (1999) constructed the “universal” type space for CPS, i.e. a type space which encodes all “conceivable” hierarchical beliefs.

Consider the following definition:

**Definition 9** A belief-complete type space is a type space $\mathcal{D}$ such that, for every $i \in N$, $g_i : \Psi_i \to \Delta^H (\Theta_0 \times \Omega_{-i})$ is onto.

Battigalli and Siniscalchi (1999) showed that a belief-complete type space may always be constructed (for all finite games, and also a large class of infinite games) by taking the sets of epistemic types to be the collection of all possible hierarchies of conditional probability systems that satisfy certain intuitive coherency conditions. Also, every type space may be viewed as a belief-closed subspace of the space of infinite hierarchies of conditional beliefs. Finally, since we assume that the set of external states is finite and hence compact, the belief-complete type space thus constructed is also compact.

**Sequential Rationality.** Fix a type space $\mathcal{D}$. For every player $i \in N$, let $f_i = (f_i,h)_{h \in \mathcal{H}} : \Omega_i \to \Delta (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i)^H$ denote his first-order belief mapping, that is, for all $\omega_i \in \Omega_i$ and $h \in \mathcal{H}$,

$$f_i, h (\omega_i) [\theta_0, \theta_{-i}, x_{-i}, s_i] = \int_{\omega \in \Omega \text{ s.t.} (\theta'_0, \theta'_{-i}, x'_{-i}, s_{-i}) = (\theta_0, \theta_{-i}, x_{-i}, s_i)} dg_{i,h}^*(\omega_i).$$

(9)

It is easy to see that $f_i$ is continuous and that, for every $\omega_i \in \Omega_i$, $f_i(\omega_i)$ is indeed a CPS over $\Theta \times X_{-i} \times S_i$, where for each $h \in \mathcal{H}$, the corresponding conditioning event $[h]_i \subseteq (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i)$ is defined as

$$[h]_i = \{(\theta_0, \theta_{-i}, x_{-i}, s_i) : (x_{-i} (\theta_{-i}), s_i) \in S (h)\}$$

(10)

For any CPS $\eta^i \in \Delta^H (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i)$, define the map $\zeta_i : \Delta^H (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) \to \Delta^H (\Theta_0 \times \Theta_{-i} \times S_i)$ so that, for each $\eta \in \Delta^H (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i)$, $\zeta_i (\eta) = (\zeta_i, h (\eta))_{h \in \mathcal{H}}$ and for each $h \in \mathcal{H}$ and $[\theta_0, \theta_{-i}, s_{-i}, s_i] \in \Theta_0 \times \Theta_{-i} \times S$,

$$\zeta_{i,h} (\eta) [\theta_0, \theta_{-i}, s_{-i}, s_i] = \sum_{x'_{-i} : x'_{-i} (\theta_{-i}) = s_{-i}} \eta (h) [\theta_0, \theta_{-i}, x'_{-i}, s_i]$$

(11)

In words, $\zeta (\eta)$ represents the payoff relevant component of $\eta$, obtained by ignoring the information on counterfactual beliefs encoded by $(x^\omega_{-i} (\theta_{-i}))_{\theta_{-i} \neq \theta'_{-i}}$. It’s easy to verify that $\zeta_i (\eta)$ thus defined is a CPS whenever $\eta$ is.
Finally, we can define rationality: we say that player $i$ is rational at a state $\omega$ if and only if $\left( \theta_i^\omega, x_i^\omega (\theta_i^\omega) \right) \in r_i (\zeta_i (f_i (\omega_i)))$. Then the event

$$\text{Rat}_i = \{ \omega \in \Omega : \left( \theta_i^\omega, x_i^\omega (\theta_i^\omega) \right) \in r_i (\zeta_i (f_i (\omega_i))) \}$$

(12)

corresponds to the statement “player $i$ is rational”. Define also the event “everybody is rational”:

$R = \bigcap_{i \in N} R_i$.

Beliefs Operators. The next building block is the epistemic notion of (conditional) probability one belief (or certainty). Recall that an epistemic type encodes the beliefs a player would hold, should any one of the possible non-terminal histories occur. This allows us to formalize statements such as, “Player $i$ would be certain that Player $j$ is rational, were he to observe history $h$”.

Given a type space $\mathcal{D}$, for every $i \in N$, $h \in \mathcal{H}$ and $E \in \mathcal{A}$, define the event

$$B_{i,h} (E) = \{ \omega \in \Omega : g_{i,h}^\omega (\omega_i) [E] = 1 \}$$

which corresponds to the statement “Player $i$ would be certain of $E$, were he to observe history $h$”. Observe that this definition incorporates the natural requirement that a player only be certain of events which are consistent with her own (continuation) strategy and epistemic type (recall how $g_{i}^\omega$ was defined, equation 8). Recalling that $h^0$ is the empty history, $B_{i,h^0} (E)$ is the event “Player $i$ believes $E$ at the beginning of the game”.

For each player $i \in I$ and history $h \in \mathcal{H}$, the definition identifies a set-to-set operator $B_{i,h} : \mathcal{A} \rightarrow \mathcal{A}$ which satisfies the usual properties of falsifiable beliefs (see, for example, Chapter 3 of Fagin et al. (1995)); in particular, it satisfies:

- **Conjunction**: For all events $E, F \in \mathcal{A}$, $B_{i,h} (E \cap F) = B_{i,h} (E) \cap B_{i,h} (F)$;
- **Monotonicity**: For all events $E, F \in \mathcal{A}$, $E \subseteq F$ implies $B_{i,h} (E) \subseteq B_{i,h} (F)$.

The mutual belief operator is defined as $B_h (E) = \bigcap_{i \in N} B_{i,h} (E)$. We will be interested in iterated beliefs operators. In general, fix a self-map on $\mathcal{A}$, $\Phi : \mathcal{A} \rightarrow \mathcal{A}$. For any $E \in \mathcal{A}$, let $\Phi^1 (E) = \Phi (E)$, and for $k = 2, 3, \ldots$, define $\Phi^k (E) = \Phi (\Phi^{k-1} (E))$. The event “common belief in $E$ at $h$” is thus defined as $\bigcap_{k \geq 1} B_h^k (E)$.

### 6.3 Common Certainty of Future Rationality

For every $h$, every $i$ and $\mu_i \in \Delta^\mathcal{H} (\Theta_0 \times \Theta_{-i} \times S)$, let $r_i (\mu_i | h)$ denote the set of pairs payoff type-strategy that are sequential best response to CPS $\mu_i$ from history $h$ onwards.
That is: for any $\mu^i \in \Delta^H(\Theta \times \Theta_{-i} \times S)$,

$$(\theta_i, s_i) \in r_i (\mu^i | h) \text{ if and only if}$$

$$\forall h' \succeq h, \quad s_i \in \arg \max_{s'_i \in S_i} \tilde{U}_i (s'_i, \mu^i; h'; \theta_i).$$

For every $i$ and $h \in \mathcal{H}$, the event “$i$ is rational in the continuation from $h$” is defined as follows:

$$\text{Rat}_{i,h} \{ \omega \in \Omega : (\theta^\omega_i, x^\omega_i (\theta^\omega_i)) \in r_i (f_i (\omega_i)) | h) \}, \quad (13)$$

The event rationality in the continuation from $h$ is defined as $\text{Rat}_h = \bigcap_{i \in N} \text{Rat}_{i,h}$.

(Clearly, $\text{Rat}_h = \text{Rat}_{h_0}$)

The event common certainty of future rationality (CCFR) corresponds to the assumption that, at each point in the game, there is common certainty that everybody is rational in the continuation game. It is formally defined as:

$$\text{CCFR} = \bigcap_{h \in \mathcal{H}} \left( \bigcap_{k \geq 1} B^k_h (\text{Rat}_h) \right).$$

The next proposition shows that $B\mathcal{R}$ corresponds to the epistemic assumptions of rationality and common certainty of future rationality.

**Proposition 3** $B\mathcal{R}_i = \text{proj}_{\Theta_i \times S_i} \text{Rat} \cap \text{CCFR}$

**Proof.** (See Appendix )

### 6.4 Belief Persistence

We introduce here an alternative characterization of $B\mathcal{R}$.

**Definition 10** For each $i \in N$ let $B\mathcal{P}_i^0 = \Theta_i \times S_i$. Recursively, for $k = 1, 2, \ldots$, and $\theta_i \in \Theta_i$, let $B\mathcal{P}_{-i}^{k-1} = \times_{j \in N \setminus \{i\}} B\mathcal{P}_j^{k-1}$,

$$B\mathcal{P}_i^k (\theta_i) = \begin{cases} \exists \eta \in \Delta^H(\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) \text{ s.t.} \\
(A) \tilde{s}_i \in r_i (\zeta^i (\eta), \theta_i) \\
(B) \text{supp} (\text{marg}_{X_{-i} \times S_i} \eta (h^0)) \subseteq B\mathcal{P}_{-i}^{k-1} \times \{ \tilde{s}_i \} \\
(C) \text{for each } h \in \mathcal{H}, \text{ supp} \left( \text{marg}_{X_i \times S_i} \eta^i (h) \right) = \text{supp} \left( \text{marg}_{X_{-i} \times S_i} \eta^i (h^0) \right) \end{cases}$$

$$B\mathcal{P}_i^k = \{ (\theta_i, s_i) \in \Theta_i \times S_i : s_i \in B\mathcal{P}_i^k (\theta_i) \}, \quad B\mathcal{P}_i^k = \times_{i \in N} B\mathcal{P}_i^k, \text{ and finally } B\mathcal{P} := \bigcap_{k \geq 0} B\mathcal{P}_i^k.$$

\(^{14}\)For a definition of $\tilde{U}_i (s_i, \mu^i; h', \theta_i)$, see equation 1, p. 8.
Condition (C) can be interpreted as a belief persistence hypothesis: at each history (whether it’s reached with positive probability or not), the support of the conjectures about the continuation strategies (one’s own and the opponents’) must not change. Condition (C) thus conforms to what in the philosophy literature is known as the conservativity principle. Such principle states that ‘When changing beliefs in response to new evidence, you should continue to believe as many of the old beliefs as possible’ (Harman, 1986, p. 46). Here, agent $i$’s conjectures over $X^h_i$ represent his “theory” of everyone’s (future) behavior in the game, both in the states $(\theta_0, \theta_{-i})$ in the support of $\eta^i$ and those that are not (similarly, conjectures over $S^h_i$ regard $i$’s beliefs about his own continuation). Definition 11 requires that upon observing history $h$, agent $i$ revises his beliefs so to accommodate the new information but without changing his “theory” about everyone’s behavior in the continuation game. Notice though that no restriction are imposed on the agent’s beliefs about the opponents’ private information, i.e. on $\text{marg}_{\Theta_{-i},\mu^i}(h)$. For example, it may concentrate beliefs on a payoff type profile $\theta_{-i}$ for which the observed history is irrational, but the restriction on the beliefs on the continuation strategies entails that it is believed that type will behave rationally in the future. In this sense, belief persistence here only refers to the endogenous variables (strategies), not the exogenous ones (payoff-types).

**Proposition 4** For each $i$, and $k$, $\mathcal{BR}^k = \mathcal{BP}^k$.

**Proof.** (See Appendix) ■

6.4.1 Common Certainty of “full” Rationality and Belief Persistence.

The event “player $i$ is rational”, defined in equation 12, consists of the set of states of the world $\omega$ in which $i$’s actual behavior $x^\omega_i (\theta^*_i)$ is a (sequential) best response to $i$’s beliefs at that state. No restrictions are imposed on the counterfactual behavior, $x^\omega_i (\theta'_i)$ for $\theta'_i \neq \theta^*_i$. We introduce next a stronger notion of rationality, which also restricts $i$’s counterfactual behavior.

$$\text{Rat}^*_i = \{ \omega \in R_i : \forall \theta'_i \in \Theta_i, (\theta'_i, x^\omega_i (\theta'_i)) \in \text{proj}_{\Theta_i \times \mathcal{S}_i} \text{Rat}_i \}$$  \hspace{1cm} (14)

$\text{Rat}^*_i$ can thus be interpreted as the set of states of the world in which $i$ not only is rational, but also his disposition to act in the (counterfactual) hypothesis that he observed a different payoff-type is consistent with rationality. The event $\text{Rat}^*_i$ thus describes states of the world in which $i$’s ex-ante strategy is rational. $\text{Rat}^*_i$ is referred to as the event “$i$ is fully rational”. As above, the event “everyone is fully rational” is defined as $\text{Rat}^* = \bigcap_{i \in N} \text{Rat}^*_i$.

---

15The conservativity principle, or the principle of belief persistence, is one of the guiding principles adopted in the philosophy literature on theories of rational belief changes (see, e.g., Gardenfors, 1988). Battigalli and Bonanno (1991) provide a semantic and syntactic characterization of the principle.
Definition 11 Say that $\Delta$-restrictions satisfy the belief persistence hypothesis if:

$$\Delta_{\theta_i}^{BP} = \{ \eta \in \Delta^H(\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) : \forall h \in \mathcal{H},$$

$$\supp \left( \text{marg}_{X_{-i} \times S_i}^h(\eta(h)) \right) = \supp \left( \text{marg}_{X_{-i} \times S_i}^h(\eta(h)) \right) \}.$$ 

For given collection $\Delta = ((\Delta_{\theta_i})_{\theta_i \in \Theta_i})_{i \in N}$ where $\Delta_{\theta_i} \subseteq \Delta^H(\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i)$, define the event

$$[\Delta_{\theta_i}] = \{(\theta_0, \theta_i, x_i, \psi_i, \omega_{-i}) \in \Omega : f_i(\theta_i, x_i, \psi_i) \in \Delta_{\theta_i}\},$$

$$[\Delta_i] = \bigcap_{\theta_i \in \Theta_i} [\Delta_{\theta_i}] \text{ and } [\Delta] = \bigcap_{i \in N} [\Delta_i]$$

Proposition 5 For any belief-complete type space:

1. $\mathcal{BR}_i^1 = \text{proj}_{\Theta_i \times S_i, \text{Rat}} \cap [\Delta^{BP}]$

2. for every $k \geq 1$, $\mathcal{BR}_i^{k+1} = \text{proj}_{\Theta_i \times S_i, \text{Rat}} \cap [\Delta^{BP}] \cap \bigcap_{v=1}^k B_{h_0}^v (\text{Rat}^* \cap [\Delta^{BP}] );$

3. if the type space is also compact, then $\mathcal{BR}_i = \text{proj}_{\Theta_i \times S_i, \text{Rat}} \cap [\Delta^{BP}] \cap \bigcap_{k=1} B_{h_0}^k (\text{Rat}^* \cap [\Delta^{BP}] ).$

7 Further Remarks on the Solution Concepts.

Backwards procedure, Subgame-Perfect Equilibrium and IPE. In games with complete and perfect information, the “backwards procedure” $\mathcal{R}(h^0)$ coincides with the backward induction solution, hence with subgame perfection.\footnote{For the special case of games with complete information, Perea (2009) independently introduced a procedure that is equivalent to $\mathcal{R}^\theta$, and showed that $\mathcal{R}^\theta$ coincides with the backward induction solution if the game has perfect information.} The next example (borrowed from Perea, 2009) shows that if the game has complete but imperfect information, strategies played in Subgame-Perfect Equilibrium (SPE) may be a strict subset of $\mathcal{R}(h^0)$:

Example 1

Consider the game in the following figure:
\[ R_1(h^0) = \{bc, bd, ac\} \quad \text{and} \quad R_2(h^0) = \{f, g\}. \] The game though has only one SPE, in which player 1 chooses \( b \): in the proper subgame, the only Nash equilibrium entails the mixed (continuation) strategies \( \frac{1}{2}c + \frac{1}{2}d \) and \( \frac{3}{4}f + \frac{1}{4}g \), yielding a continuation payoff of \( \frac{11}{4} \) for player 1. Hence, player 1 chooses \( b \) at the first node.\( \square \)

In games with complete information, IPE coincides with SPE, but \( R(h^0) \) in general is weaker than subgame perfection. At first glance, this may appear in contradiction with propositions 1 and 2, which imply that \( R(h^0) \) characterizes the set of strategies played in IPE across models of beliefs. The reason is that even if the environment has no payoff uncertainty (\( \Theta \) is a singleton), the complete information model in which \( B_i \) is a singleton for every \( i \) is not the only possible: models with redundant types may exist, for which IPE strategies differ from the SPE-strategies played in the complete information model. The source of the discrepancy is analogous to the one between Nash equilibrium and subjective correlated equilibrium (Aumann, 1974). We illustrate the point constructing a model of beliefs and an IPE in which strategy \( (ac) \) is played with positive probability by some type of player 1.\(^{17}\)

Let payoffs be the same as in example 1, and consider the model \( \mathcal{T} \) such that \( T_1 = \{t_1^{bc}, t_1^{bd}, t_1^{ac}\} \) and \( T_2 = \{t_2^f, t_2^g\} \), with the following beliefs:

\[
\tau_1(t_1) \left[ t_2^f \right] = \begin{cases} 
1 & \text{if } t_1 = t_1^{bc}, t_1^{ac} \\
0 & \text{otherwise}
\end{cases}
\]

and

\[
\tau_2(t_2^g) \left[ t_1^{ad} \right] = 1, \quad \tau_2 \left( t_2^f \right) \left[ t_1^{bc} \right] = 1
\]

The equilibrium strategy profile \( \sigma \) is such that \( \forall i, \forall t_i, b_i(t_i^{si}) = s_i \). The system of beliefs agrees with the model’s beliefs at the initial history, hence the beliefs of types \( t_2^g \) and \( t_1^{ac} \)

\(^{17}\)It is easy to see that such difference is not merely due to the possibility of zero-probability types. Also the relaxation of the common prior assumption is not crucial.
are uniquely determined by Bayesian updating. For types $t_i \neq t_2, t_{ac}$, it is sufficient to set $p_i (t_i, a_i) = \tau_i (t_i)$ (i.e. maintain whatever the beliefs at the beginning of the game were) Then, it is easy to verify that $(\theta, p)$ is an IPE.

On the other hand, if $|\Theta| = 1$ and the game has perfect information (no stage with simultaneous moves), then $R (h^0)$ coincides with the set of SPE-strategies. Hence, in environments with no payoff uncertainty and with perfect information, only SPE-strategies are played in IPE for any model of beliefs.

8 Concluding Remarks.

On the Solution Concepts. Proposition 1 can be seen as the dynamic counterpart of Brandenburger and Dekel’s (1987) characterization of correlated equilibrium. The weakness of IPE (relative to other notions of perfect Bayesian equilibrium) is key to that result: the heart of proposition 2 is $BR$’s property of continuation-game consistency, which allows to analyze continuation games “in isolation”, in analogy with the logic of backward induction. The CCFR and UIA assumptions (p. 9) provide the epistemic underpinnings of the argument. To understand this point, it is instructive to compare $BR$ with Battigalli and Sinicalschi’s (2007) weak and strong versions of extensive form rationalizability (EFR), which correspond respectively to the epistemic assumptions of (initial) common certainty of rationality (CCR) and common strong belief in rationality (CSBR): $BR$ is stronger than the first, and weaker than the latter. The strong version of EFR fails the property of “subgame consistency” because it is based on a forward induction logic, which inherently precludes the possibility of envisioning continuations “in isolation”: by taking into account the possibility of counterfactual moves, agents may draw inferences from their opponents’ past moves and refine their conjectures on the behavior in the continuation. The weak version of EFR fails continuation-game consistency for opposite reasons: an agent can make weaker predictions on the opponents’ behavior in the continuation than he would make if he envisioned the continuation game “in isolation”, because no restrictions on the agents’ beliefs about their opponents’ rationality are imposed after an unexpected history. Thus, the form of backward induction reasoning implicit in IPE (which generalizes the idea of subgame perfection) is based on stronger (respectively, weaker) epistemic assumptions than CCR (respectively, CSBR).

Appendix
A  Proofs

A.1  Proof of Proposition 1.

Step 1: \((\iff)\). Fix type space \(T = (T_i, \theta_i, \tau_i)_{i \in N}\), IPE \((\hat{b}, \hat{p})\), type \(\hat{t}_i\) such that \(\theta_i(\hat{t}_i) = \hat{\theta}_i\), and \(\hat{s}_i\). For each \(j\), and \(\theta_j \in \Theta_j\), let

\[
\bar{S}_j(\theta_j) = \left\{ s_j \in S_j : \exists t_j \in T_j \text{ s.t. } \theta_j(t_j) = \theta_j \text{ and } s_j \in \text{supp} \left( \hat{b}_j(t_j) \right) \right\},
\]

\[
\bar{S}_j = \{ (t_j, s_j) : s_j \in \bar{S}_j(\theta_j(t_j)) \}
\]

and \(S^h_j = \{ (t_j, s^h_j) \in T_j \times S^h_j : \exists s_j \in S_j(\theta_j(t_j)) \text{ s.t. } s_j|h = s_j \}\).

We will prove that \(\bar{S}_j(\theta_j) \subseteq \mathcal{B} R_j(\theta_j)\) for every \(j\) and \(\theta_j\). For each \(h^l_i = (\hat{t}_i, h^l) \in \mathcal{H}_i\) and \(t_j \in T_j\), define \(\varphi^j_{t_j} : \bar{S}_j(\theta_j(t_j)) \rightarrow S_j(h^l)\) such that

\[
\varphi^j_{t_j}(s_j)(h^r) = \left\{ \begin{array}{ll}
  s_j(h^r) & \text{if } \tau > l \\
  a^r_j & \text{otherwise}
\end{array} \right.
\]

where \(a^r_j\) is the action played by \(j\) at period \(\tau \leq l\) in the public history \(h^l\). Thus, \(\varphi^j_{t_j}\) transforms any interim strategy in \(\bar{S}_j\) into one that has the same continuation from \(h^l\), and that agrees with the public history \(h^l\) for the previous periods. Define the mapping \(L^l_i : \Theta_0 \times T_{-i} \times S \rightarrow \Theta_0 \times T_{-i} \times S(h^l)\) such that

\[
\forall (\theta_0, t_{-i}, s_i, s_{-i}) \text{ s.t. } s_i \in \bar{S}_i(\theta_i(\hat{t}_i)) \times \bar{S}_{-i}(\theta_{-i}(t_{-i})),
\]

\[
L^l_i(\theta_0, t_{-i}, s_i, s_{-i}) = (\theta_0, t_{-i}, \varphi^j_{t_j}(s_i), \varphi^j_{t_{-i}}(s_{-i})).
\]

In particular, at the interim stage \(h^0_i = (\hat{t}_i, h^0)\), \(L^0_i(\theta_0, t_{-i}, s) = (\theta, t_{-i}, s)\).

For each \(h^l_i = (\hat{t}_i, h^l)\), let \(P^{\hat{b}, \hat{p}}(h^l_i) \in \Delta(\Theta_0 \times T_{-i} \times S)\) denote the probability distribution on \(\Theta_0 \times T_{-i} \times S_{-i}\) induced by \(\hat{p}_i(h^l_i)\) and \(\hat{b}\):

\[
P^{\hat{b}, \hat{p}}(\theta_0, t_{-i}, s_i, s_{-i}; h^l_i) = p_1(\theta_0, t_{-i}; h^l_i) \cdot \hat{b}_i(s_i; \hat{t}_i) \cdot \hat{b}_{-i}(s_{-i}; t_{-i})\]

Define the CPS \(\lambda_i \in \Delta^{H_i}(\Theta_0 \times T_{-i} \times S)\) as follows:

for \(h^0_i = (\hat{t}_i, h^0)\), let

\[
\lambda_i(\hat{t}_i, h^0) = \lambda_i(h^0_i) = \left\{ \begin{array}{c}
\lambda_i(h^0_i) \left[ (\theta_0, t_{-i}, s) = P^{\hat{b}, \hat{p}}(\theta_0, t_{-i}, s; h^l_i) \right] = 0 \\
\text{and for all } h^l_i = (\hat{t}_i, h^{l-1}, a^l) \in \mathcal{H}_i \text{ such that } \lambda_i(\hat{t}_i, h^{l-1}) \left[ \{(\theta_0, t_{-i}, s) : s(h^{l-1}) = a^l \} \right] = 0, \text{ let } \\
\lambda_i(h^l_i) = \sum_{(\theta_0, t_{-i}, s') \in \Theta_0 \times T_{-i} \times S : (\theta_0, t_{-i}, s') = L^l_i(\theta_0, t_{-i}, s')} P^{\hat{b}, \hat{p}}(\theta_0, t_{-i}, s'; h^l_i) .
\end{array} \right.
\]
(conditional beliefs \(\lambda_i(h_i^t)\) at histories \(h_i^t\) s.t. \(\lambda_i(h_i^{t-1})\left[\{(\theta_0, t_{-i}, s) : s \left(h_i^{t-1} = a^t\right)\}\right] > 0\) are determined via Bayesian updating, from the definition of CPS, Section 6.1)

Define the CPS \(\mu^i \in \Delta^{\mathcal{H}_i} (\Theta_0 \times \Theta_{-i} \times S)\) s.t. \(\forall h_i \in \mathcal{H}_i\),

\[
\mu^i(h_i^t)[\theta_0, \theta_{-i}, s] = \sum_{t_{-i} \theta_{-i}(t_{-i}) = \theta_{-i}} \lambda_i(h_i^t)[\theta_0, t_{-i}, s].
\]

By construction, \(\hat{b}_i(\hat{t}_i) \in r_i(\mu^i)\). We only need to show that conditions (2) and (3) in the definition of BR are satisfied by \(\mu^i\). This part proceeds by induction: The initial step, for \(k = 1\), is trivial. Hence, \(\tilde{S}_j \subseteq \mathcal{BR}_j\) for every \(j\). To complete the proof, let (as inductive hypothesis) \(\tilde{S}_j \subseteq \mathcal{BR}_j\) so that condition (2) in \(\mathcal{BR}_j\) are satisfied by \(\mu^j\). We can construct a map \(\hat{b}_i(\hat{t}_i)\) such that \(\mu^i(\hat{t}_i, h_i^0) \subseteq \Theta^* \times \{\hat{b}_i(\hat{t}_i)\} \times \mathcal{BR}_j\) and

\[
\text{supp}(\text{marg}_{\tilde{S}_j \mu^i}(\phi)) = \text{supp}(\text{marg}_{\tilde{S}_j \mu^i}(h_i^0)) \subseteq \tilde{S}_j^{h_i^0}.
\]

Thus \(\hat{b}_i(\hat{t}_i) \in \mathcal{BR}_i^{k+1}\). This concludes the first part of the proof.

**Step 2:** \((\Rightarrow)\). Let \(T\) be such that for each \(i\), \(T_i = \mathcal{BR}_i\), \(\theta_i : T_i \rightarrow \Theta_i\) such that \(\forall (\theta_i, s_i) \in T_i\), \((\theta_i, s_i) = \theta_i\) and let strategy \(\hat{b}_i : T_i \rightarrow S_i\) be such that \(\forall (\theta_i, s_i) \in T_i\), \(\hat{b}_i(\theta_i, s_i) = s_i\). Let \(t_i : T_i \rightarrow \Theta_i \times S_i\) denote the identity map.

Notice that, for each \(i, \theta_i\) and \(s_i \in \mathcal{BR}_i(\theta_i), \exists \mu^{s_i} \in \Delta^{\mathcal{H}} (\Theta_0 \times \Theta_{-i} \times S)\) s.t.

1. \(s_i \in r_i(\mu^{s_i}, \theta_i)\)
2. for all \(h \in \mathcal{H}\): \((\theta_j, s_j) \in \text{supp}(\text{marg}_{\Theta_j \times S_j \mu^{s_j}}(h)) \Rightarrow \exists s'_j \in \mathcal{BR}_j(\theta_j) : s_j|h^{t-1} = s'_j|h^{t-1}.

Hence, for each \(h\), we can construct a map \(\rho_{s_i,h} : \text{supp}(\text{marg}_{\Theta_{-i} \times S_{-i} \mu^{s_i}}(h)) \rightarrow \mathcal{BR}_{-i}\) such that

\[
\forall (\theta_{-i}, s_{-i}) \in \text{supp}(\text{marg}_{\Theta_{-i} \times S_{-i} \mu^{s_i}}(h)), \rho_{s_i,h}(\theta_{-i}, s_{-i}) = (\theta_{-i}, s'_{-i}) \text{ s.t. } s'_{-i} \in \mathcal{BR}(\theta_{-i}) \text{ and } s'_{-i}|h = s_{-i}|h.
\]

Define the map \(M_{i,0} : \Theta_0 \times \mathcal{BR}_{-i} \times S_i \rightarrow \Theta_0 \times \Theta_{-i}\) s.t.

\[
M_{i,0}(\theta_0, \theta_{-i}, s_i) = \left(\theta_0, t_{-i}^{-1}(\theta_{-i}, s_{-i})\right).
\]

Let \(m_{s_i,h} \equiv t_{-i}^{-1} \circ \rho_{s_i,h}\). Define maps \(M_{s_i,h} : \text{supp}(\mu^{s_i}(h)) \rightarrow \Theta_0 \times \Theta_{-i}\)

\[
M_{s_i,h}(\theta_0, \theta_{-i}, s_{-i}) = (\theta_0, m_{s_i,h}(\theta_{-i}, s_{-i})).
\]
Let beliefs \( \tau_i : T_i \to \Delta (\Theta_0 \times T_{-i}) \) be s.t. \( \forall (\theta_0, t_{-i}) \in \Theta_0 \times T_{-i} \)

\[
\tau_i(t_i)[\theta_0, t_{-i}] = \sum_{(\theta_0', \theta_{-i}', s') : (\theta_0, t_{-i})=M_{i,0}(\theta_0, \theta_{-i}, s)} \mu^{\hat{b}(t_i)}(\phi)[\theta_0', \theta_{-i}', s']
\]

Let beliefs \( \hat{\pi}_i \) be derived from \( \hat{b} \) and the initial beliefs \( \hat{\pi}_i(t_i, h^0) = \tau_i(t_i) \) via Bayesian updating whenever possible. At all other information sets \( h_i = (t_i, h) \in \mathcal{H} \), for every \( (\theta_0, t_{-i}) \in \Theta_0 \times T_{-i} \), set

\[
\hat{\pi}_i(\theta_0, t_{-i}|h_i) = \sum_{(\theta_0', \theta_{-i}', s') : (\theta_0, t_{-i})=M_{i,h}(\theta_0, \theta_{-i}, s)} \mu^{b(t_i)}(h)[\theta_0', \theta_{-i}', s']
\]

By construction, \((\hat{\sigma}, \hat{\pi})\) is an IPE of \( \Gamma^T \) \( \blacksquare \)

### A.2 Proof of Proposition 2

**Step 1** \((R_i(h^0) \subseteq \mathcal{BR}_i)\): let \( \hat{s}_i \in R_i(\theta_i, h^0) \). Then, for each \( h \in \mathcal{H} \), \( s_i|h \in R_i(\theta_i, h) \). Notice that for each \( h \) and \( s^h_j \in R_i(\theta_i, h) \), there exists \( s_i \in R_i(\theta_i, h^0) \) such that \( s_i|h = s^h_i \).

Thus, for each \( j, \theta_j \) and \( h \), we can define mappings \( \rho^h_j : R_j(\theta_j, h) \to R_j(\theta_j, h^0) \) such that:

\( \forall s^h_j \in R_j(\theta_j, h) \)

\[
\rho^h_j(s^h_j) | h = s^h_{j/}
\]

(Functions \( \rho^h_j \) assign to strategies in \( R_j(\theta_j, h) \), strategies in \( R_j(\theta_j, h^0) \) with the same continuation from \( h \).) As usual, denote by \( \rho^h_{-j} \) the product \( \times_{j \neq i} \rho^h_j \).

For each \( h^l \), let \( \varphi^h_{j^l} : S_j \to S_j(h^l) \) be such that

\[
\varphi^h_{j^l}(s_j)(h^\tau) = \begin{cases} s_j(h^\tau) & \text{if } \tau > l \\ a^\tau_j & \text{otherwise} \end{cases}
\]

where \( a^\tau_j \) is the action played by \( j \) at period \( \tau \leq t \) in the public history \( h^l \). (As usual, denote by \( \varphi^h_{-j} \) the product \( \times_{j \neq i} \varphi^h_j \).

For each \( h^l \), define the mapping \( \varphi^h_{i, -i} : R_{-i}(h^l) \to \Theta_{-i} \times S_{-i}(h^l) \) such that \( \forall (\theta_{-i}, s^h_{-i}) \in R_{-i}(h^l) \),

\[
\varphi^h_{i, -i}(\theta_{-i}, s^h_{-i}) = (\theta_{-i}, \varphi^h_{-i} \circ \rho^h_{-i}(s^h_{-i})).
\]

It will be shown that: for each \( k = 0, 1, ..., \hat{s}_i \in R^k_i(\theta_i, h^0) \) implies \( \hat{s}_i \in \mathcal{BR}_i^k(\theta_i) \).

The initial step is trivially satisfied \((\mathcal{BR}_i^0(\theta_i) = S_i = R^0_i(\theta_i, h^0))\).

For the inductive step, suppose that the statement is true for \( n = 0, ..., k - 1 \): Since \( \hat{s}_i \in R^k_i(\theta_i, h^0) \), for each \( h^l \) there exists \( \pi^{h^l} \in \Delta (\Theta_0 \times R_{-i}(h^l)) \) that satisfy

\[
\hat{s}_i|h^l \in \arg \max_{s'_i|h^l} \sum_{(\theta_0, \theta_{-i}, s^h_{-i}) \in \Theta_0 \times R_{-i}(h^l)} U_i(s_i, s_{-i}, \theta_i, \theta_{-i}; h^{l-1}) \cdot \pi^{h^l}[\theta_0, \theta_{-i}, s^h_{-i}].
\]
Now, consider the CPS \( \mu^i \in \Delta^H (\Theta_0 \times \Theta_{-i} \times S) \) such that, for all \((\theta_0, \theta_{-i}, s_i, s_{-i}) \in \Theta \times S, \)

\[
\mu^i (\theta_0, \theta_{-i}, s_i, s_{-i}|h^0) = \begin{cases} 
\pi^{h_0} \left[ \theta_0, \theta_{-i}, s_{-i}^h \right] & \text{if } s_i = \hat{s}_i \\
0 & \text{otherwise}
\end{cases}
\]

By definition of CPS, \( \mu^i (h^0) \) defines \( \mu (h^l) \) for all \( h^l \) reached with positive probability. Let \( h^l = (h^{l-1}, a^l) \in H \) be such that \( \mu^i (h^0) [h^{l-1}] > 0 \) and \( \mu^i (h^0) [h^l] = 0 \). Define \( M_{h^l, \hat{s}_i} : \Theta_0 \times R_{-i} (h^l) \rightarrow \Theta_0 \times \Theta_{-i} \times S (h^l) \) such that for all \((\theta_0, \theta_{-i}, s_{-i}^h) \in \Theta_0 \times R_{-i} (h^l), \)

\[
M_{h^l, \hat{s}_i} (\theta_0, \theta_{-i}, s_{-i}^h) = \left( \theta_0, \varphi_i^h (\hat{s}_i) \vartheta_i^h \left( \theta_{-i}, s_{-i}^h \right) \right)
\]

and set \( \mu^i (h^l) \) equal to the pushforward of \( \pi^{h^l} \) under \( M_{h^l, \hat{s}_i} \), i.e. \( \forall (\theta_0, \theta_{-i}, s) \in \Theta_0 \times \Theta_{-i} \times S (h^l) \)

\[
\mu^i (\theta_0, \theta_{-i}, s|h^l) = \sum_{(\theta'_0, \vartheta_{-i}, s_{-i}^h) \in \Theta_0 \times R_{-i} (h^l): M_{h^l, \hat{s}_i} (\theta'_0, \vartheta_{-i}, s_{-i}^h) = (\theta_0, \theta_{-i}, s)} \pi^{h^l} \left[ \theta'_0, \vartheta_{-i}, s_{-i}^h \right].
\]

Again, by definition of CPS, \( \mu^i (h^l) \) defines \( \mu (h^r) \) for all \( h^r \succ h^l \) that receive positive probability under \( \mu^i (h^l) \). For other histories, proceed as above, setting \( \mu^i (h^r) \) equal to the pushforward of \( \pi^{h^r} \) under \( M_{h^r, \hat{s}_i} \), and so on.

By construction, \( \hat{s}_i \in r_i (\mu^i, \theta_i) \) (condition 1 in the definition of \( \mathcal{BR}_i \)). Since by construction \( \mu^i (h^0) [\Theta_0 \times R_{-i}^{h_{-i}^0} (h^0) \times \{\hat{s}_i\}] = 1 \), under the inductive hypothesis \( \mu^i (h^0) [\Theta_0 \times R_{-i}^{h_{-i}^{k-1}} \times \{\hat{s}_i\}] = 1 \) (condition 2 in the definition of \( \mathcal{BR}_i \)). From the definition of \( \varphi_i^h (\hat{s}_i) \), CPS \( \mu^i \) satisfies condition (3.1) at each \( h^l \). From the definition of \( \vartheta_i^h \), under the inductive hypothesis, \( \mu^i \) satisfies condition (3.2).

**Step 2 (\( \mathcal{BR}_i \subseteq R_i (h^0) \)):** let \( \hat{s}_i \in R_i (\theta_i) \) and \( \mu^i \in \Delta^H (\Theta_0 \times \Theta_i \times S) \) be such that \( \hat{s}_i \in r_i (\mu^i, \theta_i). \) For each \( h^l \), define the mapping \( \psi_i^h : S_{-i} \rightarrow S_{-i}^{h^l} \) s.t. \( \psi_i^h (s_{-i}) = s_{-i} | h^l \) for each \( s_{-i} \in S_{-i}. \) (Function \( \psi_i^h \) transforms each strategy profile of the opponents into its continuation from \( h^l \).)

For each \( h^l \in H \), let \( \pi^{h^l} \in \Delta \left( \Theta_0 \times \Theta_{-i} \times S_{-i}^{h^l} \right) \) be such that for every \((\theta_0, \theta_{-i}, s_{-i}^h) \in \Theta_0 \times \Theta_{-i} \times S_{-i}^{h^l} \)

\[
\pi^{h^l} \left[ \theta_0, \theta_{-i}, s_{-i}^h \right] = \sum_{(s_{-i}, s_{-i}) \in S (h^l): \psi_i^h (s_{-i}) = s_{-i}^h} \mu^i \left( \theta_0, \theta_{-i}, s_{-i}, s_{-i}|h^l \right).
\] (15)

so that the implied joint distribution over payoff states and continuation strategies \( s_{-i}|h^l \) is the same under \( \mu^i (h^l) \) and \( \pi^{h^l} \). We will show that, for every \( h^l \in H, \hat{s}_i|h^l \in R_i (\theta_i, h^l). \)
Notice that, by construction,
\[ \hat{s}_i|h^l \in \arg \max_{s_i \in S_i^h} \sum_{(\theta,\theta_{-i},s_{-i}) \in \Theta \times \Theta_{-i} \times S_{-i}^h} U_i(s_i,s_{-i},\theta_i,\theta_{-i};h^l) \cdot \pi^{h^l}[\theta_0,\theta_{-i},s_{-i}]. \]

The argument proceeds by induction on the length of histories.

**Initial Step** ($L=1$). Fix history $h^{L-1}$: we shall prove that for each $k$, if $\hat{s}_i \in BR_i^k(\theta_i)$, then $\hat{s}_i|h^{L-1} \in R_i^k(\theta_i,h^{L-1})$.

For $k = 0$, it is trivial. For the inductive step, let $\pi^{h^{L-1}}$ be defined as above: under the inductive hypothesis, $\pi^{h^{L-1}}(\Theta_0 \times R_{-i}^{k-1}(h^{L-1})) = 1$ (condition R.1), while (16) implies that condition R.2 is satisfied.

**Inductive Step:** suppose that for each $\tau = l + 1,\ldots,L$, $\hat{s}_i \in BR_i(\theta_i)$, implies $\hat{s}_i|h^\tau \in R_i(\theta_i,h^\tau)$ for each $h^\tau \in H^\tau$. We will show that for each $h^l \in H^l$, for each $k = 0,1,\ldots,\hat{s}_i|h^l \in R_i^k(\theta_i,h^l)$. We proceed by induction on $k$: under the inductive hypothesis on $\tau$, $\hat{s}_i|h^l \in R_i^0(\theta_i,h^l)$. For the inductive step on $k$, suppose that $\hat{s}_i \in BR_i(\theta_i)$, implies $\hat{s}_i|h^l \in R_i^0(\theta_i,h^l)$ for $n = 0,\ldots,k-1$, and suppose (as contrapositive) that $\hat{s}_i|h^l \notin R_i^k(\theta_i,h^l)$. Then, for $\pi^h$ defined as in (15) (which ensures (16)), it must be that $\text{supp}(\pi^h) \not\subseteq \Theta_0 \times R_{-i}^{k-1}(h^l)$, which, under the inductive hypothesis on $n$, implies that $\exists (\theta_{-i},s_{-i}) \in \text{supp}(\text{marg}_{\Theta_{-i} \times S_{-i}}\mu^i(h^l))$ s.t. $\exists s'_{-i} \in BR_{-i}(\theta_{-i}) : s'_{-i}|h^l = s_{-i}|h^l$, which contradicts that $\mu^i$ justifies $\hat{s}_i$ in $BR_i$.

**B Epistemic Characterizations**

**B.1 Proof of Proposition 4**

The proof is by induction. For $k = 0$, the statement is trivially satisfied. Now, suppose (as inductive hypothesis) that $BR_j^m = BP_j^m$ for all $m = 0,1,\ldots,k-1$, we will show that $BR_i^k = BP_i^k$.

**Step 1:** ($BP_i^k \subseteq BR_i^k$) Let $(s_1,\eta^i)$ satisfy conditions (A-C) in definition 10, and let $\mu^i = \zeta_i(\eta^i)$. Under the inductive hypothesis, conditions (1) and (2) in definition 3 are clearly satisfied by $(s_i,\mu^i)$. Furthermore, given (A) and (B), if $\eta$ satisfies (C) then $\zeta_i(\eta^i)$ satisfies (C).

**Step 2:** ($BR_i^k \subseteq BP_i^k$) Let $(s_1,\mu^i)$ satisfy conditions (1-3) in definition 3. For each $h$, let $\chi^h : \text{supp} \text{marg}_{\Theta_{-i} \times S_{-i}}\mu^i(h) \rightarrow X_{-i}$ be such that for each $\hat{\theta}_{-i},\hat{s}_{-i} \in \Theta_{-i} \times S_{-i}$, $\chi^0(\hat{\theta}_{-i},\hat{s}_{-i}) = \hat{x}_{-i}$ satisfies:

(a) $\hat{x}_{-i}(\hat{\theta}_{-i}) = \hat{s}_{-i},$
(b) $\hat{x}_{-i}(\theta_{-i}) \in BR_{-i}^{k-1}(\theta_{-i})$ for all $\theta_{-i},$ and
(c) $\hat{x}_{-i}(\theta_{-i})|h = \hat{x}_{-i}(\hat{\theta}_{-i}) = \hat{s}_{-i}|h$
(Conditions (b) and (c) can be satisfied for any \( \left( \tilde{\theta}_{-i}, \tilde{s}_{-i} \right) \in \text{supp } \text{marg}_{\Theta_{-i} \times S_{-i}} \mu^i (h) \) because \( \mu^i \) satisfied condition (3) in definition 3.

It is convenient to introduce the following piece of notation: \( \eta \in [\Delta (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i)]^H \), and for any \( h^l, h^{l+k} \in \mathcal{H} \), say that

\[
h^{l+k} \in \text{supp } \eta (h^l)
\]

if and only if

\[
\exists s \in \left[ \left\{ S \left( h^{l+k} \right) \cap \{(s_{-i}, s_i) : \exists (\theta_0, \theta_{-i}, x_{-i}, s_i) \in \text{supp}_x \left( h^l \right) \text{ s.t. } x_{-i} (\theta_{-i}) = s_{-i} \} \right\} \right]
\]

In words, \( h^{l+k} \in \text{supp } \eta (h^l) \) means conditional conjectures \( \eta (h^l) \) assign positive probability to \( h^{l+k} \) being reached.

Now, let \( \eta^0 \in \Delta^H (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) \) be such that, for all \( (\theta_0, \theta_{-i}, x_{-i}, s_i) \in \Theta_0 \times \Theta_{-i} \times X_{-i} \times S_{-i} \),

\[
\eta^0 (h^0) \left[ \theta_0, \theta_{-i}, x_{-i}, s_i \right] = \begin{cases} \mu^i (h^0) \left[ \theta_0, \theta_{-i}, s_{-i}, s_i \right] & \text{if } x_{-i} = \chi^{h^0} (\theta_{-i}, s_{-i}) \\ 0 & \text{otherwise} \end{cases}
\]

and for all \( h \in \text{supp } \eta^0 (h^0) \), let \( \eta^0 (h) \) be obtained by Bayesian updating.

Now, let \( E^0 = \{ h^0 \} \), and recursively, for \( \nu = 1, 2, \ldots \), let \( E^\nu \subseteq \mathcal{H} \):

\[
E^\nu = \left\{ h^l \in H : \exists h^{l-1} \in E^{\nu-1} \text{ s.t. } h^l \notin \text{supp } \eta^0 \left( h^{l-1} \right) \text{ and } h^l = (h^{l-1}, a') \text{ for some } a' \in A^l \right\}.
\]

For each \( \hat{h} \in E^\nu \), let \( \eta^0 \left( \hat{h} \right) \) be such that, for all \( (\theta_0, \theta_{-i}, x_{-i}, s_i) \in \Theta_0 \times \Theta_{-i} \times X_{-i} \times S_{-i} \),

\[
\eta^0 \left( \hat{h} \right) \left[ \theta_0, \theta_{-i}, x_{-i}, s_i \right] = \begin{cases} \mu^i \left( \hat{h} \right) \left[ \theta_0, \theta_{-i}, s_{-i}, s_i \right] & \text{if } x_{-i} = \chi^{\hat{h}} (\theta_{-i}, s_{-i}) \\ 0 & \text{otherwise} \end{cases}
\]

and for all \( h' \in \text{supp } \eta^0 \left( \hat{h} \right) \), let \( \eta^0 \left( h' \right) \) be obtained by Bayesian updating. By construction, \( \mu^i = \zeta_i \left( \eta^0 \right) \), hence condition (A) in definition 10 is satisfied. Condition (C) is automatically satisfied at all histories \( h \in \text{supp} \eta^0 (h^0) \). For all others, it is sufficient to modify \( \eta^0 (h^0) \) so that the continuations of the \( x_{-i} \) in its support at such histories \( h \notin \text{supp} \eta^0 (h^0) \), so that they agree with those in the support of \( \eta^0 (h) \). (the idea is simple, but notationally very involved). Notice that this changes don’t affect the distributions on the payoff relevant variables. That is, calling \( \eta \) the CPS thus obtained, \( \zeta_i \left( \eta \right) = \zeta_i \left( \eta^0 \right) \), hence by construction \( \eta \) satisfies conditions (A) (B) and (C).

**B.2 Proof of Proposition 3**

The proof of proposition 3 exploits an auxiliary solution concept based on the notion of weak \( \Delta \)-rationalizability introduced by Battigalli and Siniscalchi (2007).
The Auxiliary Solution Concept (1). The auxiliary solution concept in this case in nothing but Battigalli and Siniscalchi’s (2007) weak $\Delta$-rationalizability applied to the continuation game starting from some public history $h$. For each history $h$, it’s useful to define the set of histories in the continuation game from $h$, i.e. $H(h) = \{h' \in H : h' \succeq h\}$. Clearly, $h$ is the initial node in this continuation game. The auxiliary solution concept specifies a correspondence $W_{i,h}^\Delta : \Theta_i \rightrightarrows S_i^h$ for each player $i$, taking as given a profile $\Delta^h$ of information-dependent first-order restrictions: formally, let $\Delta^h = ((\Delta^h_{0_i})_{\theta_i \in \Theta_i})_{i \in N}$ where $\Delta^h_{0_i} \subseteq \Delta^{H(h)}(\Theta_0 \times \Theta_{-i} \times S_{-i}^h)$ is a nonempty closed set for each payoff type $\theta_i$ of each player $i$. Then, $\Delta$-weak rationalizability in the continuation game is defined as follows:

**Definition 12** For each $\theta_i \in \Theta_i$, $W_{i,h}^\Delta(\theta_i) = \cap_{k \geq 0} W_{i,h}^{\Delta k}(\theta_i)$, where $W_{i,h}^{\Delta 0}(\theta_i) = S_i^h$ and, recursively, $W_{i,h}^{\Delta k}(\theta_i)$ is the set of all continuation strategies $s_i \in S_i^h$ such that (a)

\[
   s_i \in r_i(\mu, |h\|) \text{ for some } \mu \in \Delta^h_{0_i} \text{ that satisfies:} \ (17)
\]

\[
   \text{supp} \left( \text{marg}_{\theta_{-i} \times S_{-i}^h}^i(\theta_i) \right) \subseteq W_{-i,h}^{\Delta k-1} \times \{s_i\} \ . \ (18)
\]

For given collection $\Delta^h = ((\Delta^h_{0_i})_{\theta_i \in \Theta_i})_{i \in N}$, define the event

\[
   [\Delta^h_{0_i}] = \{ (\theta_0, \theta_i, x_i, \psi_i, \omega_{-i}) \in \Omega : (f_{i,h'} (\theta_i, x_i, \psi_i))_{h' \in H(h)} \in \Delta_{\theta_i} \},
\]

\[
   [\Delta^h] = \bigcap_{\theta_i \in \Theta_i} [\Delta^h_{0_i}] \text{ and } [\Delta_h] = \bigcap_{i \in N} [\Delta^h_i]
\]

The following result is follows from minor adaptations of proposition 1 in Battigalli and Siniscalchi (2007, BS).

**Proposition 6** For any belief-complete type space:

1. $W_{i,h}^{\Delta 1} = \text{proj}_{\Theta_i \times S_i^h} \text{Rat}_h \cap [\Delta_h]$

2. for every $k \geq 2$, $W_{i,h}^{\Delta k} = \text{proj}_{\Theta_i \times S_i^h} \left( \text{Rat}_h \cap [\Delta_h] \cap \bigcap_{\nu=1}^{k-1} B_h^\nu (\text{Rat}_h \cap [\Delta_h]) \right)$;

3. if the type space is also compact, then $W_{i,h}^{\Delta} = \text{proj}_{\Theta_i \times S_i^h} \bigcap_{k \geq 1} B_h^k (\text{Rat}_h \cap [\Delta_h])$.

**B.2.1 Proof of Proposition 3**

The proof of proposition 3 exploits the characterization of $\text{BR}$ in terms of the backwards procedure presented in Section 5. The proof is by induction, starting at the end of the game:
Initial Step (l = L − 1): for each h ∈ H^{L−1},

\[(\theta_i, s_i) \in \text{proj}_{\Theta_i \times S_i} \text{Rat}_h \cap \left( \bigcap_{k \geq 1} B^k_h (\text{Rat}_h) \right) \]

if and only if

\[(\theta_i, s_i|h) \in R_i (h) \]

The initial step follows from proposition 6, once it’s proven that \( R_i (h) = W_{i,h}^\Delta \), letting the \( \Delta \)-restrictions be the trivial restrictions, i.e. such that \( \Delta^h_{\theta_i} = \Delta \left( \Theta_0 \times \Theta_{−i} \times S^h_{−i} \right) \) for each \( \theta_i \). This is immediate though, given that the continuation game is static at \( h \in H^{L−1} \).

Inductive Step (l = L − 2, ..., 0): Suppose that (19) holds for all histories \( h \in H^\nu \), where \( \nu = l + 1, ..., L − 1 \). Then, for each \( h \in H^l \),

\[(\theta_i, s_i) \in \text{proj}_{\Theta_i \times S_i} \text{Rat}_h \cap \left( \bigcap_{h' \in \mathcal{H} (h)} \bigcap_{k \geq 1} B^k_{h'} (\text{Rat}_{h'}) \right) \]

if and only if

\[(\theta_i, s_i|h) \in R_i (h) \]

Poof of the Inductive Step: the inductive step follows from 6, once it’s proven that \( R_i (h) = W_{i,h}^\Delta \) where the \( \Delta \)-restrictions are such that, for each \( i \) and \( \theta_i \),

\[\Delta^h_{\theta_i} = \left\{ \mu \in \Delta (\mathcal{H} (h)) \left( \Theta_0 \times \Theta_{−i} \times S^h_{−i} \right) : \forall h' \in \mathcal{H} (h), \right\} \]

if \( (\theta_0, \theta_{−i}, s_{−i}) \in \text{supp} \mu \), then:

\[\forall h^{l+1} = (h, a_{−i}^{l+1}) \in H^{l+1}, (\theta_{−i}, s_{−i}|h^{l+1}) \in W_{i,h^{l+1}} \].

Suppose that we have proven that we have proven that \( R_i (h') = W_{i,h'}^\Delta \) for all \( h' \in H^\nu \) s.t. \( \nu = l + 1, ..., L − 1 \).

- [\( W_{i,h} \subseteq R_i (h) \)]: Under the inductive hypothesis, the \( \Delta \)-restrictions thus specified imply that

\[\{ \pi \in \Delta (\Theta_0 \times \Theta_{−i} \times S^h_{−i}) : \mu = \pi (h) \text{ for some } \mu \in \Delta^h_{\theta_i} \}\]

= \( \Delta (\Theta_0 \times R^0_{−i} (h)) \)

Furthermore, if \( (\theta_i, s_i) \in r_i (\mu|h) \) for some \( \mu \in \Delta^h_{\theta_i} \) (i.e. \( s_i \) is a sequential best reply to \( \mu \)), then \( s_i \in R^1_i (\theta_i, h) \) (this is because in \( R^1_i \) strategy \( s_i \) survives if it’s
best response at \( h \). Hence, \( W^1_{i,h} \subseteq R^1_i (h) \). Inductively, if \( W^\nu_{i,h} \subseteq R^\nu_j (h) \) for every \( \nu = 1, \ldots, k \) and \( j \), then

\[
\Delta \left( \Theta_0 \times R^k_{-i} (h) \right) \supseteq \left\{ \pi \in \Delta \left( \Theta_0 \times \Theta_{-i} \times S^h_{-i} \right) : \pi = \mu (h) \right. \\
\left. \quad \text{for some } \mu \in \Delta^h_{\theta_i} \text{ s.t. } \mu (h) \in \Delta \left( \Theta_0 \times W^k_{-i} \right) \right\}
\]

Hence \( W^{k+1}_{i,h} \subseteq R^{k+1}_i (h) \).

- \([R_i (h) \subseteq W^{\Delta}_{i,h}] : \) let \((\theta_i, s_i) \in R^1_i (h)\) and \( \hat{\pi} \in \Delta \left( \Theta_0 \times R^0_{-i} (h) \right) \) be the justifying conjecture. Since, by definition, \( R^1_i (h) \subseteq R^0_i (h) \), for every \( h' = (h, a) \) s.t. \( a \in A (h) \), \((\theta_i, s_i | h') \in R_i (h') \). Hence, under the inductive hypothesis, \((\theta_i, s_i | h') \in r_i (\eta | h')\) for some \( \eta' \in \Delta^{H(h')} \left( \Theta_0 \times W^{\Delta}_{-i,h'} \right) \). Now, let \( \mu \in \Delta^{H(h')} \left( \Theta_0 \times W^{\Delta_0}_{-i,h'} \right) \) be s.t. \( \mu (h) = \hat{\pi} \) and for all \( h' = (h, a) \) s.t. \( \mu (h) [h'] = 0 \), for all \( h'' \in H (h') \) set \( \mu (h'') = \eta' (h'') \). Then, by construction, \( s_i \in r_i (\mu | h) \), and since under the inductive hypothesis (21) holds, \((\theta_i, s_i) \in W^{\Delta_1}_{i,h} \). Now, suppose that we have shown that \( R^\nu_i (h) \subseteq W^{\Delta^\nu}_{i,h} \) for every \( \nu = 1, \ldots, k \), and let \((\theta_i, s_i) \in R^{k+1}_i (h) \), with justifying conjecture \( \hat{\pi} \in \Delta \left( \Theta_0 \times R^k_{-i} (h) \right) \) be the justifying conjecture. Since, by definition, \( R^{k+1}_i (h) \subseteq R^0_i (h) \), for every \( h' = (h, a) \) s.t. \( a \in A (h) \), \((\theta_i, s_i | h') \in R_i (h') \). Hence, under the inductive hypothesis, \((\theta_i, s_i | h') \in r_i (\eta | h')\) for some \( \eta' \in \Delta^{H(h')} \left( \Theta_0 \times W^{\Delta}_{-i,h'} \right) \). Notice that, under the inductive hypothesis,

\[
\Delta \left( \Theta_0 \times R^k_{-i} (h) \right) \subseteq \left\{ \pi \in \Delta \left( \Theta_0 \times \Theta_{-i} \times S^h_{-i} \right) : \pi = \mu (h) \right. \\
\left. \quad \text{for some } \mu \in \Delta^h_{\theta_i} \text{ s.t. } \mu (h) \in \Delta \left( \Theta_0 \times W^k_{-i} \right) \right\}
\]

Now, let \( \mu \in \Delta^{H(h')} \left( \Theta_0 \times W^{\Delta^k}_{-i,h} \right) \) be s.t. \( \mu (h) = \hat{\pi} \) and for all \( h' = (h, a) \) s.t. \( \mu (h) [h'] = 0 \), for all \( h'' \in H (h') \) set \( \mu (h'') = \eta' (h'') \). Then, by construction, \( s_i \in r_i (\mu | h) \), and since (21) holds, \((\theta_i, s_i) \in W^{\Delta_{k+1}}_{i,h} \).

### B.3 Proof of Proposition 5

The proof of proposition 5 exploits an auxiliary solution concept based on the notion of weak \( \Delta \)-rationalizability introduced by Battigalli and Siniscalchi (2007).

**The Auxiliary Solution Concept (2).** Similar to weak \( \Delta \)-rationalizability, this solution concept specifies a correspondence \( A^\Delta_i : \Theta_i \rightrightarrows S_i \) for each player \( i \), taking as given a profile \( \Delta \) of information-dependent first-order restrictions: formally, \( \Delta = [(\Delta_{\theta_i})_{\theta_i \in \Theta_i}]_{i \in N} \) where \( \Delta_{\theta_i} \subseteq \Delta^h (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) \) is a nonempty closed set for each payoff type \( \theta_i \) of each player \( i \). The auxiliary solution concept is defined as follows:
Definition 13 For each \( \theta_i \in \Theta_i \), \( A^\Delta_i(\theta_i) = \cap_{k \geq 0} A^{\Delta,k}_i(\theta_i) \), where \( A^{\Delta,0}_i(\theta_i) = S_i \) and, recursively, \( A^{\Delta,k}_i(\theta_i) \) is the set of all \( s_i \in S_i \) such that (a)

\[
s_i \in r_i \left( \zeta_i(\eta^i), \theta_i \right) \text{ for some } \eta^i \in \Delta_{\theta_i} \text{ that satisfies: (22)}
\]

\[
\text{supp} \left( \text{marg}_{X_{-i} \times S_i} \eta^i (h_0) \right) \subseteq \left\{ x_{-i} \in X_{-i} : \forall \theta_{-i} \in \Theta_{-i}, x_{-i}(\theta_{-i}) \in A_i^{\Delta,k-1}(\theta_{-i}) \right\} \times \{s(\mathcal{P}3)\}
\]

For given collection \( \Delta = ((\Delta_{\theta_i})_{\theta_i \in \Theta_i})_{i \in N} \) where \( \Delta_{\theta_i} \subseteq \Delta^H(\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) \), define the event

\[
[\Delta_{\theta_i}] = \{(\theta_0, \theta_i, x_i, \psi_i, \omega_{-i}) \in \Omega : f_i(\theta_i, x_i, \psi_i) \in \Delta_{\theta_i}\},
\]

\[
[\Delta_i] = \bigcap_{\theta_i \in \Theta_i} [\Delta_{\theta_i}] \text{ and } [\Delta] = \bigcap_{i \in N} [\Delta_i]
\]

The following result is parallels Proposition 1 in Battigalli and Siniscalchi (2007, BS).

Proposition 7 For any belief-complete type space:

1. \( A^{\Delta,1}_i = \text{proj}_{\Theta_i \times S_i} R \cap [\Delta] \)

2. For every \( k \geq 2 \), \( A^{\Delta,k}_i = \text{proj}_{\Theta_i \times S_i} \left( R \cap [\Delta] \cap \bigcap_{\nu=1}^{k-1} B^{\nu}_{h_0}(R^* \cap [\Delta]) \right) \);

3. If the type space is also compact, then \( A^\Delta_i = \text{proj}_{\Theta_i \times S_i} \bigcap_{k=0}^k B^k_{h_0}(R \cap [\Delta]). \)

The proof is based on the following lemma, which follows from minor modifications of the proof in BS. (Its proof is thus omitted here.)

Lemma 2 (cf. Lemma 6, BS) Fix \( (\hat{\theta}_i, \hat{s}_i) \in \Theta_i \times S_i \), a map \( \gamma_{-i} : X_{-i} \to \Psi_{-i} \) and first-order CPS \( \delta^i \in \Delta^H(\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) \). Then there exists \( (\hat{\theta}_i, x_i, \psi_i) \in \Omega_i \) such that \( x_i(\hat{\theta}_i) = \hat{s}_i \) and \( \forall h \in \mathcal{H} \), \( g^*_i(h, \theta_i, x_i, \psi_i) \in \Delta(\Omega) \) has finite support and \( \text{marg}_{\Theta_0 \times \Theta_{-i} \times X_{-i} \times \Psi_{-i} \times S_i} g^*_i(h, \theta_i, x_i, \psi_i) \left[ \theta_0, \theta_{-i}, x_{-i}, \gamma_{-i}(x_{-i}), s_i \right] = \delta^i(h) [\theta_0, \theta_{-i}, x_{-i}, s_i] \)

Proof of Proposition 7

- For \( k = 0 \), I show that \( A^{\Delta,1}_i = \text{proj}_{\Theta_i \times S_i} R \cap [\Delta] \):

\[
R \cap [\Delta] = \left\{ \omega = \left( \theta_0, (\omega_j)_{j \in N} \right) \in \Omega : \forall i \in N \right.
\]

\[
f_i(\omega_i) \in \Delta_{\theta_i} \text{ and } (\theta_i, s_i) \in r_i \left( \zeta_i \left( f_i(\omega_i) \right) \right) \}
\]

it is immediate to see that \( (\theta_i, s_i) \in \text{proj}_{\Theta_i \times S_i} R \cap [\Delta_i] \) implies that \( s_i \in A^{\Delta,1}_i(\theta_i) \). For the opposite direction, for each \( j \in N \) pick arbitrarily \( \gamma^0_j : X_j \to \Psi_j \). Now, for each
\[i \in I \text{ and each } (\theta_i, s_i) \in A_{i}^{\Delta_1}, \text{ let } \eta \in \Delta_{\theta_i} \text{ be s.t. } (\theta_i, s_i) \in r_i(\zeta_i(\eta)). \text{ Then, from lemma } 2, \text{ there exists } \gamma^1_i(x_i) \in \Psi_i \text{ such that for all } h \in \mathcal{H}, g^*_i h (\theta_i, x_i, \gamma^1_i(x_i)) [\theta_0, \theta_{-i}, \gamma^{0}_{-i}(x_{-i}), s_i] = \eta(h) [\theta_0, \theta_{-i}, x_{-i}, s_i], \text{ hence } \zeta_i (f_i(\theta_i, x_i, \gamma^1_i(x_i))) = \zeta_i(\eta).

Define \(\hat{\gamma}^1_i : \Theta_j \times S_j \to \Omega_j \text{ s.t.: } \forall \left(\hat{\theta}_j, \hat{s}_j\right) \in A_{j}^{\Delta_1}, \hat{\gamma}^1_j (\hat{\theta}_j, \hat{s}_j) = \left(\hat{\theta}_j, x_j, \gamma^1_j(x_j)\right)\) s.t. \(x_j(\hat{\theta}_j) = \hat{s}_j\). Then, by construction, for each \(\hat{\theta}_i, \hat{s}_i \in A_1^{\Delta_1}\), let \(\gamma_i \text{ and } \gamma^1_i(x_i) \) such that \(x_i(\hat{\theta}_i) = s_i\) (this completes the definition of \(\hat{\gamma}^1_i : X_i \to \Omega_i \text{ on the full domain}).

Before proceeding with the induction, it is useful to notice that, for any \(E \in A\) and for every \(\mathbb{m}\)

\[
\bigcap_{\mathbb{m}=1}^{m_1} B_{h^0}(E) = \bigcap_{\mathbb{m}=0}^{m_1-1} B_{n,h^0}(\Omega_{i} \cap \text{proj}_{\theta_i} B_{h^0}^{*_{n+1}}(E))
\]

- For \(k = 1 \text{ and } (\theta_i, s_i) \in A_1^{\Delta_1} \text{ and } \eta \in \Delta_{\theta_i} \text{ be the justifying CPS s.t. supp } \text{ marg}_{X_{-i} \times S_i} \eta(h^0) \subseteq A_{-1} \times \{s_i\}. \text{ Under the inductive hypothesis, for every } (\theta_0, \theta_{-i}, x_{-i}, s_i) \in \text{supp } \text{ marg}_{X_{-i}} \eta(h^0), \text{ there exists } \gamma^1_{-i}(x_{-i}) \in \Psi_{-i} \text{ s.t. } (\theta_{-i}, x_{-i}, \gamma^1_{-i}(x_{-i})) \in \text{proj}_{\theta_{-i}} (R^* \cap [\Delta]).

Then, from lemma 2, there exists \(\omega \) s.t. \(x^0 (\theta_i^0) = s_i \text{ and } \gamma^1_i(x_i) \text{ such that for all } h \in \mathcal{H}, g^*_i h (\theta_i, x_i, \gamma^1_i(x_i)) [\theta_0, \theta_{-i}, \gamma^1_{-i}(x_{-i}), s_i] = \eta(h) [\theta_0, \theta_{-i}, x_{-i}, s_i]. \text{ Hence, by construction,}\)

\[
(\theta_i, x_i, \gamma^2_i(x_i)) \in \text{proj}_{\theta_i} B_{i,h^0} (\Omega_i \cap \text{proj}_{\theta_{-i}} (R^* \cap [\Delta]))
\]

(and, since \(x_i(\theta_i) = s_i\), also \(\theta_i, x_i, \gamma^2_i(x_i) \in \text{proj}_{\theta_i} R \cap [\Delta].)\)

Define \(\hat{\gamma}^2_i : \Theta_j \times S_j \to \Omega_j \text{ s.t.: } \forall \left(\hat{\theta}_j, \hat{s}_j\right) \in A_{j}^{\Delta_2}, \hat{\gamma}^2_j (\hat{\theta}_j, \hat{s}_j) = \left(\hat{\theta}_j, x_j, \gamma^2_j(x_j)\right)\) s.t. \(x_j(\hat{\theta}_j) = \hat{s}_j\). Then, by construction, for each \(\hat{\theta}_i, \hat{s}_i \in A_1^{\Delta_2}\), \(\omega = \left(\theta_0, \left(\hat{\gamma}^2_i (\hat{\theta}_i, \hat{s}_i)\right)_{i \in N}\right) \subseteq R \cap [\Delta] \cap B^{1}_{h^0} (R^* \cap [\Delta])\). To prove the other direction, let \(\left(\theta_0, (\theta_i, x_i, \psi_i)_{i \in N}\right) \subseteq R \cap [\Delta] \cap B^{1}_{h^0} (R^* \cap [\Delta])\). Then, for each \(i\),

\[
g_i h^0 (\theta_i, x_i, \psi_i) [R \cap [\Delta] \cap B^{1}_{h^0} (R^* \cap [\Delta])] = 1.
\]

Hence, under the inductive hypothesis, it’s immediate to verify that \(x_i(\theta_i) \in r_i(\zeta_i (f_i(\theta_i, x_i, \psi_i)))\), \(f_i(\theta_i, x_i, \psi_i) \in \Delta_{\theta_i}\), and that \(f_i h^0 (\Theta_0 \times \Theta_{-i} \times A_{-1} \times \{x_i(\theta_i)\}) = 1\). Hence, \(\theta_i, x_i(\theta_i) \in A^2_{i} \Delta\).

- For \(k \geq 2\), suppose that for \(m = 1, \ldots, k-1\), \(A_{i}^{\Delta,m+1} = \text{proj}_{\theta_i \times S_i} R \cap [\Delta] \cap \bigcap_{\mathbb{m}=1}^{m} B_{h^0} (R^* \cap [\Delta])\) and that for each such \(m\) we have defined functions \(\gamma_{i}^{m+1} : X_i \to \Psi_i \) and \(\hat{\gamma}_{i}^{m+1} : \Theta_i \times S_i \to \Omega_i \) as above, such that \(\left(\theta_0, (\hat{\gamma}^{m+1}_i (\theta_i, s_i))_{i \in N \in N}\right) \subseteq R \cap [\Delta] \cap \bigcap_{\mathbb{m}=1}^{m} B_{h^0}^{\ast_{n+1}} (R^* \cap [\Delta])\) whenever \(\theta_i, s_i \in A^m_{i} \Delta.\)
- \( A^\Delta, k+1 \subseteq \text{proj}_{\Theta \times S} R \cap [\Delta] \cap \bigcap_{i=1}^{k} B^\nu_0 \) (\( R^* \cap [\Delta] \))

Let \( (\theta_i, s_i) \in A^\Delta, k+1 \) and \( \eta \in \Delta_{\Theta_i} \), be the justifying CPS s.t. \( \text{supp marg}_{X \times \Theta_i \times S_i} \eta (h^0) \subseteq A_{\Theta_i} \times \{s_i\} \). Under the inductive hypothesis, for every \( (\theta_0, \theta_i, x_i, s_i) \in \text{supp marg}_{X \times \Theta_i \eta (h^0)} \), there exists \( \gamma^k_{x-i} (x_i) \in \Psi_i \) s.t. \( (\theta_i, x_i, \gamma^k_{x-i} (x_i)) \in \text{proj}_{\Theta_i \cap [\Delta] \cap \bigcap_{i=1}^{k} B^\nu_i \} (R^* \cap [\Delta]) \). Then, from lemma 2, there exists \( \omega \) s.t. \( x^\omega_i (\theta^*_{i}) = s_i \) and \( \gamma_{i}^{k+1} (x_i) \) such that for all \( h \in \mathcal{H}, g_i^* (\theta_i, x_i, \gamma_i^{k+1} (x_i)) \left[ \theta_0, \theta_i, \gamma_i^{k+1} (x_i), s_i \right] = \eta (h) [\theta_0, \theta_i, x_i, s_i] \).

Hence, by construction,

\[
(\theta_i, x_i, \gamma_i^{k+1} (x_i)) \in \text{proj}_{\Theta_i \times S_i} \left( \Omega_i \cap \text{proj}_{\Theta_i} \left( \bigcap_{i=1}^{k} B^\nu_i \cap [\Delta] \right) \right) \\
(\text{and, since } x_i (\theta_i) = s_i \text{, also } (\theta_i, x_i, \gamma_i^{k+1} (x_i)) \in \text{proj}_{\Theta_i \cap [\Delta] \cap \bigcap_{i=1}^{k} B^\nu_i \} (R^* \cap [\Delta]) \).
\]

Define \( \hat{\gamma}^{k+1}_j : \Theta_j \times S_j \rightarrow \Omega_j \) s.t.: \( \forall \left( \hat{\theta}_j, \hat{s}_j \right) \in A^\Delta, j+1, \hat{\gamma}^{k+1}_j \left( \hat{\theta}_j, \hat{s}_j \right) = \left( \hat{\theta}_j, x_j, \gamma_j^{k+1} (x_j) \right) \) s.t. \( x_j (\hat{\theta}_j) = \hat{s}_j \). Then, by construction, for each \( \hat{\theta}_i, \hat{s}_i \in \hat{\Omega}_i \), \( \omega = \left( \hat{\theta}_0, \left( \hat{\gamma}^{k+1}_i \left( \hat{\theta}_i, \hat{s}_i \right) \right) \right) \in R \cap [\Delta] \cap \bigcap_{i=1}^{k} B^\nu_i \) (\( R^* \cap [\Delta] \)).

\[
\text{proj}_{\Theta_i \times S} R \cap [\Delta] \cap \bigcap_{i=1}^{k} B^\nu_i \) (\( R^* \cap [\Delta] \)) \subseteq A^\Delta, k+1
\]

Let \( (\theta_0, (\theta_i, x_i, s_i)) \in R \cap [\Delta] \cap \bigcap_{i=1}^{k} B^\nu_i \) (\( R^* \cap [\Delta] \)). Then, for each \( i, g^*_i (\theta_i, x_i, s_i) \left[ R \cap [\Delta] \cap \bigcap_{i=1}^{k-1} B^\nu_i \) (\( R^* \cap [\Delta] \left) = 1. \right. \) Hence, under the inductive hypothesis, it's immediate to verify that \( x_i (\theta_i) \in r_i (\zeta_i (f_i (\theta_i, x_i, s_i))) \), \( f_i (\theta_i, x_i, s_i) \in \Delta_{\Theta_i} \), and that \( f_i \left( \theta_i, x_i, s_i \right) \in \Delta_{\Theta_i} \). Finally, suppose that \( A^\Delta, j+1 = \mathcal{B} \mathcal{P}_j \) for all \( j \in N \) and \( m = 0, ..., k-1 \). Then, under the inductive hypothesis, \( \eta \) satisfies conditions (B) in definition 10 if and only if it satisfies equation (23).

**B.3.1 Proof of Proposition 5**

The proof of Proposition 5 is based on an equivalence result between \( \mathcal{B} \mathcal{P} \) and \( A^\Delta \), provided that the \( \Delta \)-restrictions satisfy the **belief persistent hypothesis**. That is (definition 11: \( \forall i \in N, \forall \theta_i \in \Theta_i \))

\[
\Delta^\mathcal{B \mathcal{P}}_i = \{ \eta \in \Delta^N (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) : \forall h \in \mathcal{H}, \text{supp} \left( \text{marg}_{X \times \Theta_{-i} \times S_i} \eta (h) \right) = \text{supp} \left( \text{marg}_{X \times \Theta_{-i} \times S_i} \eta (h^0) \right) \}.
\]

**Lemma 3** If the \( \Delta \)-restrictions satisfy the belief persistence hypothesis, then for every \( \theta_i \in \Theta_i \) and for every \( k = 0, 1, ..., A^\Delta, k (\theta_i) = \mathcal{B} \mathcal{P}^k (\theta_i) \).

**Proof.** The proof is by induction. The initial step \( (k = 0) \) is trivially satisfied; for \( k = 1, 2, ..., \) notice that \( \eta \in \Delta^N (\Theta_0 \times \Theta_{-i} \times X_{-i} \times S_i) \) satisfies condition (C) in definition 10 if and only if \( \eta \in \Delta^\mathcal{B \mathcal{P}}_i \). Finally, suppose that \( A^\Delta, \eta = \mathcal{B} \mathcal{P}^m \) for all \( j \in N \) and \( m = 0, ..., k-1 \); then, under the inductive hypothesis, \( \eta \) satisfies conditions (B) in definition 10 if and only if it satisfies equation (23).
Proposition 5 follows immediately from Lemma 3 and proposition 7.

References


